

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 22:10:58 ; Search time 876.5 Seconds

(without alignments)
6590.438 Million cell updates/sec

Title: US-10-664-859-1

Perfect score: 12037
Sequence: 1 acgagtgctctcttatat.....aaaaaaaaaaaaaaaaaaaaa 6909

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p,model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10664859/runat_02112005_171806_3577/app_query.fasta_1.7047
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cst -LIST=45 -LOCALIGN=200 -THR SCORE=cct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10664859@cgn_1_1_982@runat_02112005_171806_3577
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
17: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
18: /cgn2_6/prodata/2/pubppaa/US10C_PUBCOMB.pep:*
19: /cgn2_6/prodata/2/pubppaa/US11_PUBCOMB.pep:*
20: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7040	58.5	1429	20	US-11-097-143-3129
2	570	4.7	112	9	US-09-915-543-10
3	570	4.7	112	14	US-10-322-579-10
4	570	4.7	112	15	US-10-664-859-10
5	321.5	2.7	1435	15	US-10-276-774-2178
6	314	2.6	1394	15	US-10-381-247B-17
7	314	2.6	1394	16	US-10-723-860-1764
8	314	2.6	1394	18	US-10-756-149-5163
9	312	2.6	1426	9	US-09-915-543-15
10	312	2.6	1426	14	US-10-322-579-15
11	312	2.6	1426	15	US-10-664-859-15
12	312	2.6	1426	16	US-10-408-765A-1958
13	296.5	2.5	1450	15	US-10-381-247B-17
14	296.5	2.5	1494	15	US-10-381-247B-2
15	269.5	2.2	1594	20	US-11-097-143-17856
16	258.5	2.1	1366	20	US-11-097-143-27555
17	258.5	2.1	2151	20	US-11-097-143-7050
18	256.5	2.1	5322	17	US-10-732-923-8729
19	247.5	2.1	2285	18	US-10-773-446-101
20	242.5	2.0	1161	20	US-11-097-143-38793
21	241.5	2.0	1778	20	US-11-097-143-7035
22	240	2.0	2280	20	US-11-097-143-11742
23	233.5	1.9	1233	16	US-10-408-765A-1964
24	233.5	1.9	1441	15	US-10-355-218-2
25	233.5	1.9	1441	18	US-10-503-050A-2
26	233.5	1.9	2703	20	US-11-097-143-16689
27	233	1.9	3190	17	US-10-732-923-18448
28	233	1.9	3275	18	US-10-840-060-90
29	230.5	1.9	3275	20	US-11-097-143-38103
30	230.5	1.9	3275	17	US-10-732-923-18447
31	230	1.9	3276	17	US-10-732-923-18447
32	229.5	1.9	2005	9	US-09-735-167B-3
33	228.5	1.9	2053	9	US-09-735-167B-2
34	228.5	1.9	2063	18	US-10-973-858-8
35	228.5	1.9	2063	18	US-10-450-765-47506
36	228	1.9	1424	15	US-10-418-027-1
37	228	1.9	1424	16	US-10-473-127-347
38	228	1.9	1424	16	US-10-473-127-350
39	228	1.9	1424	16	US-10-473-127-360
40	228	1.9	1522	16	US-10-473-127-355
41	222.5	1.8	1441	14	US-10-170-682-3
42	222.5	1.8	1459	15	US-10-369-493-6418
43	222.5	1.8	1966	16	US-10-408-765A-2217
44	222.5	1.8	2440	15	US-10-341-434-236
45	222.5	1.8	2440	18	US-10-885-977-35

ALIGNMENTS

RESULT 1
US-11-097-143-3129
; Sequence 3129, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097, 143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/457, 832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160, 191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161, 932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164, 769

1 PRIOR FILING DATE: 1999-11-12
1 PRIOR APPLICATION NUMBER: 60/173,383
1 PRIOR FILING DATE: 1999-12-28
1 PRIOR APPLICATION NUMBER: 60/175,693
1 PRIOR FILING DATE: 2000-01-12
1 PRIOR APPLICATION NUMBER: 60/184,831
1 PRIOR FILING DATE: 2000-02-24
1 PRIOR APPLICATION NUMBER: 60/191,637
1 PRIOR FILING DATE: 2000-03-23
1 NUMBER OF SEQ ID NOS: 43008
1 SOFTWARE: FastSeq for Windows Version 4.0
1 SEQ ID NO 3129
1 LENGTH: 1429
1 TYPE: PRT
1 ORGANISM: DROSOPHILA
1 US-11-097-143-3129

Alignment Scores:

Pred. No.:	0	Length:	1429
Score:	7040.00	Matches:	1425
Percent Similarity:	71.09%	Conservative:	1
Best Local Similarity:	71.04%	Mismatches:	3
Query Match:	58.49%	Indels:	577
DB:	20	Gaps:	6

US-10-664-859-1 (1-6909) x US-11-097-143-3129 (1-1429)

QY 453 ATGCTCTGACACATGCCCCGAGTCCAAACCAACAGCGGCAACCAACTCCGAT 512
1 MetLeuSerThrThrMetProHisSerProThrGlnGlnInProGlnProHisSerAsp 20
QY 513 GCCTCTCAACAAATGTCATCTGATCAAAATCTGAGACAGCGATCGAAATGGGGAATCG 572
21 AlaSerSerThrSerAlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSer 40
QY 573 GCGGCGAGGAGAGTCTCTCGAAGACCTTAATAGGAAACCTTTTCTACTTTGTGCGCG 632
41 AlaAlaSerThrSerSerProLysThrLeuAsnSerGlnProHisSerThrLeu-Ser-- 59
QY 633 GGTAAAGACTTGATGATTTCTCTTGTCCGGAATTAAACAATTCTGTGTTCCAGA 692
60 -----ProAs 61
QY 693 TCAATTAATAATTGACCGCAGAGAGAGGACTGAGAAAAGCGACTATCAACTAGATTA 752
61 pGlnIleLysLeuThrProGlnGlnGlyThrGlnLysSerGlyLeuSerThrSerAsp 81
QY 753 AGCTGCACTGAGAGAGCCCGAGAGTGAATTAATCTGCCGAGAGCAAACTATGCT 812
81 sAlaAlaThrGlyAlaProGlySerGlyAsnAsnLeuProGlnGlyGlnThrMetLe 101
QY 813 AAGCAGAACTCTACAGACACATCAACTGTCCTTACAGTTCACAAAACCTCCAG 872
101 uAGGlnAsnSerThrSerThrIleAsnSerGlybLeuValAlaSerProGlnAsnSerSe 121
QY 873 TGAACACTGGAATGACAGCAATGTGTCTGCTACAGTGGGCTTACTCAAGTGTAGATTG 932
121 rGlnHisSerAsnSerSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCy 141
QY 933 TGAAGGACATCGAAGAAAACAATGTAGTGAAGAGACAGAGAGCTCGTAAAGCTGC 992
141 sAspGlnGlnSerLysLysAsnLysCySerValLysAspGlnGlnAla----- 157
QY 993 CCTACAAATGTTTAAATTTTAAATGTATGGCGCTTACCTTTGTTATCATTTAATT 1052
157 ----- 157
QY 1053 GTTTTTTTTTTCTATACTTACAAATTTTAACTTTAACTGTAACCTGTAATAAATCG 1112
157 ----- 157
QY 1113 CGAAGCTCGATCAAAAAGACATTTTCTTGGAACCGTATTTAAGCTCATTAATAATTTA 1172

Db 157 ----- 157
QY 1172 ATTCACTTGATGATGATGATATCATATGATGACTCAACATCTCAAGAAAGACCTCAAA 1232
Db 157 ----- 157
QY 1233 TTGGATCAACTAATTAGTTTGAGAAAAAATGCTGCTACTTTAAGAAATTAATTAATTAA 1292
Db 157 ----- 157
QY 1293 AAATTTGCTGAGTGAATGATATTAATAGTACAAATATTTTATGTTAAACTGTAAAGCA 1352
Db 157 ----- 157
QY 1353 TTTGTAATAGCCGTGCTACGACAGATGCTACTAGACCGCGGTGTAAGACTAATTTTATTT 1412
Db 157 ----- 157
QY 1413 AAAAGCTGCTTAATATTCATTAACCATTAATGTCCTCATTTGCAAAATTAAGTTCTAATA 1472
Db 158 -----GlutLeuSerSerAsnLys 163
QY 1473 AGCAAAAGGTCAAGCAGCTGTGTGGCGGCTGCGAAAACAGGTTCTACATCCAGTTTGACTGT 1532
Db 163 sAlaLysGlyGlnAlaAlaGlyGlyGlyCysGlnThrGlySerThrSerLeuThrVal 183
QY 1533 CAAGAAAGAACCCACGATGCTTTAGGACGTTTAAATATGAAATTAAGAAAGAAAGAGA 1592
183 LysGlnGlnProThrThrAspValLeuGlySerLeuValAsnMetLysbGlnGlnLysArg 203
QY 1593 AAATCAATTGCGCAAGATGCTCCCTGTGGTTGGTTGGTTGTTGTTGTTGTTGTTGTTGTTG 1652
203 uAsnHisSerProThrMetSerProValGlybPheGlySerIleGlyAsnAlaGlnAspAs 223
QY 1653 CTCGCTACAGCGGGTAAGTTTAAAGATCCATATAAGCAATTAACAAGATTAATGT 1712
223 nSerAlaThr----- 226
QY 1713 CAGTTACCAATTTTATTTGATGATGCAAGAACTACTATAGCGATATCTGCTGCTTTTAA 1772
226 ----- 226
QY 1773 TTTTATTTTATTAAGAAATACGAATATTTTCTAATTTGTAATAATAATGATTAATTAA 1832
226 ----- 226
QY 1833 CTAGAAATTAATAAACCTTTGAATTAGACATACCTTCCAAAAATCAGTAATCATTTGG 1892
226 ----- 226
QY 1893 AACGAGATGTGTCCGAGAGAGACTACTATMAAACCTTTGAGCTATCTGATACCTGCA 1952
226 ----- 226
QY 1953 CCCTACTAAAAATGATTAAGTAAAGAAAAATGGGTGAATTTTGTAGAAAGTTTCATTTT 2012
226 ----- 226
QY 2013 AGAAGAAATGTGATTATTTATTAACCCCTCAAGCGGACTACATTTGTTCTACGATA 2072
226 ----- 226
QY 2073 TTTTGGAAAAACMAATGTTAAGTTGAAGAAAGTGCCTATMAAACAGAAATTCACGGTTTCA 2132
226 ----- 226
QY 2133 AATTAACAGAGTTTGTGATTTAATTTGATTAATGAGAAATTAATACACTTCAGTTA 2192
226 ----- 226
QY 2193 AAATGTTTAATTGATTAAGTTCGACATCAAGAGATTTTCATTTTTCGCGTATAT 2252
226 ----- 226

```
QY 2253 ATGAGAGTCGCTTCACTCTTCTGGGGGCTTCACTACGTCAGTGGAGTTCCGGCCGA 2312
Db 226 ----- 226
QY 2313 GTGATTATATAGATGATTTACAGTATTTTATTTTATGTGTATTTTAAATATAT 2372
Db 226 ----- 226
QY 2373 CTATTTTATTCATTTTACATAGTTAAATTTGAAAGATTTTCAACGACGATCCACGGAA 2432
Db 227 -----Pro--ValysIleGluArgIleSerAsnAspSerThrThGlu 240
QY 2433 AAAAAGATCGTCCTTGACAAATGAAATGACAAATGACAAATGACAAAGGTCGCAATCAG 2492
Db 241 LysIleGlySerSerLeuThrMetAsnAsnAspIleSerMetGluGlyCyAsnGln 260
QY 2493 TTGAATCCGATTTTATCAATGAAATCTTTAAATATCTGCAATTTTGACATATTAGTA 2552
Db 261 LeuAsnProAspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuVal 280
QY 2553 AGGGAGTAGAGCAATACCCGGAATGGAGTTGGAGCGGGAGCGGAATTTATTTACT 2612
Db 281 SerGlyValGlyProIleProGlyIleGlyValGlyValGlyThrGlyAsnLeuThr 300
QY 2613 GCCAAGCCCAATGGAATCTCTCGGGTAGAGATATGTTGATTCATGCAACAGCAA 2672
Db 301 AlaAsnAlaAsnGlyIleSerSerGlySerSerAsnGlyLeuAspGlyMetGlnGlnGln 320
QY 2673 AATCACATATTCGTGTTTCACTAGCTGAGCCCAACAAAGGGCCGAATCAGTTTAAAC 2732
Db 321 AsnHisIlePheValPheSerThrGlnLeuAlaAsnGlyValGlyIleGluSerValLeuSer 340
QY 2733 GGTCAATTTTAACTATTTATGCGGTATCACTGCACTAGCCCTGTACAAAAGTTTCTG 2792
Db 341 GlyIlePheGlnThrIleIleAlaGlyHisCysThrGlnProAlaThrIleSerPheLeu 360
QY 2793 GAAGACTTTTAAAGAAAACCTTTAAAGATTAACAGTTTAAAGGCGGCAATTCGCTC 2852
Db 361 GluAspPhePheMetCysAsnProLeuLysIleAsnLysLeuGlnAlaGlnHisSerVal 380
QY 2853 GGTATGCGCATGATAGGATGAGGAGGAGTTGCACTTAATCTCTTAATCTGTAGCCAA 2912
Db 381 GlyMetProThrIleGlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLys 400
QY 2913 ATAACACACAGCAGCCACATACAAAGACGTAAGGCTTATTTAAACCCCAATTCATCAA 2972
Db 401 IleThrGlnGlnGlnProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGln 420
QY 2973 CATGAAAACGCAAGCTAGTACTGTAAAGGCGCTAGCAACTTTTGTGCAACGACT 3032
Db 421 HisGluAsnSerLysAspSerThrValSerAlaProSerAsnSerPheValAspGlnSer 440
QY 3033 GATCTATAGGCAACGAACTGAATGTATGCTGTGGAGAGCGGATCTTAAACACGACT 3092
Db 441 AspProMetClyAsnGluThrGluLeuMetCysThrGlnGlyGlySerSerAsnThrSer 460
QY 3093 AGGTCTGGACAAAACCTACGAATCATGTAGACAGTATCATGATCATCAGAGTCAAG 3152
Db 461 ArgSerGlyGlnAsnSerArgAsnHisValAspSerIleSerThrSerSerGluSerGln 480
QY 3153 GCATATAAGTACTGGAGAGAGCTGGGCTGATTTGGGACAGGCTCAAAAAGAGAGAT 3212
Db 481 AlaIleLysIleLeuGlnAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAsp 500
QY 3213 CCTGGCTGCAACCTGAAAACAATTTGATCTACTGCAAGAGATTAAAGTTCCAGACAA 3272
Db 501 ProGlyLeuThrThrGluAsnAsnIleValSerLeuGlnGlyValLysValProAspGlu 520
QY 3273 AACCTTACACCAACAGCGGCAATCGGAGAAAGACAGTTGGCAAAAATTAATAAATG 3332
Db 521 AsnLeuThrProGlnGlnArgGlnHisArgGlyGlnGlnLeuAlaLysIleLysLysMet 540
QY 3333 AATCAATTTCTTTTCTGAAAAATGAGAAATTCAGTAGAGCTATGTAGCTACAGATA 3392
Db 541 AsnGlnPheLeuPheProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIle 560
QY 3393 ACAAAAATTCAGAGATTTTAAATGATGGGATGTCCGGTGGCGGAGCGGATCTATTATA 3452
Db 561 ThrLysIleProGlyAspLeuMetMetGlyMetSerGlyGlyGlyGlySerIleIle 580
QY 3453 AATCCGACGATGGGACAACTGCAATAGCCAGGTACGCCCAATTCGGAGCTTTACGGCG 3512
Db 581 AsnProThrMetArgGlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAla 600
QY 3513 ACAAGTTCAAGACTTTTCGAAGAATGATATCATCTCAGAGGATGTTATATCAGATAGGCT 3572
Db 601 ThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMetCly 620
QY 3573 GCCGTAATAGATGTAATAATAATCAAAAAACAGTGTGCAATGTGGATTTGGATAGCT 3632
Db 621 AlaValIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGly 640
QY 3633 GTTGTCACTGGAAACAATGAGCTGAGCTGAATGTAATGTAATGCTCAAGCTCCGGC 3692
Db 641 ValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGly 660
QY 3693 GCCCGAATGCAATATGATGGAGAGCTCTACGGAATATGCTTGTGCAACACA 3752
Db 661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
QY 3753 AGCTGCAACGTATCGGAACGCGCCCAAGATATGTCCTTAAGAGTTTAAATCAAGATAGC 3812
Db 681 SerCysAsnValIleGlyThrAlaProAspMetSerGlyValLeuAsnGlnAspSer 700
QY 3813 CGAACCATTCACATCAAGGGAGAGTGTCTCAATAGAGAGTGTGCAAGATTTCAACATCAA 3872
Db 701 ArgThrHisSerHisGlnGlyGlyValAlaGlnMetGlyThrPheLysIleGlnHisGln 720
QY 3873 TTTTTCGAAGACGCTTCAGAGGGGAGGAGCCCAAGTCACTGGAACCTGATAGTACCA 3932
Db 721 PhePheGluGluArgLeuLysGlyLysProArgGlnValThrGlyThrValValPro 740
QY 3933 CAACAGCAAAACCCCTTCTGATCTGTGTAACCTGTTAAACAACAGGTGCAACCCCTG 3992
Db 741 GlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAsnGlnValArgProLeu 760
QY 3993 CAAGGTCCACTCTCTCTTACCACTGCATCAGAGATCTCCGTACAGTACCAATAGCACT 4052
Db 761 GlnGlyProProProProGlyHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
QY 4053 CAATGGCCCAATCCCTCGAAGTCCAAACAATCTATCTCTCCGTCACCGCGGACAAACGCA 4112
Db 781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProAlaGlnThrAla 800
QY 4113 GCAGTCAATGGATTGCGCAACAATCTCTTACAGATGATGAAACAGATCATTAATCTGGA 4172
Db 801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
QY 4173 TCTGTTCCGCACTAATACTTGGACGTTCAAGAGGACCAACAACAGTCTCTACGA 4232
Db 821 SerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrValIleSerAla 840
QY 4233 AACAGAATCTGTTTACAGGACAGACCCCATCTCCCGTCAATCAATCAATCTAGTAGAAAT 4292
Db 841 AsnLysAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsn 860
QY 4293 ACCGATCGTCAAGCGTTCTTACGATTAATTTAAGACAGCAACCAAGTACCCCTTAATCT 4352
Db 861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
QY 4353 CATCTATCCCAAAAGAAATTTAGTCTTTTCGGTCAAGTCCCTGCTGTATGTTATATTGG 4412
Db 881 HisLeuSerProLysGluPheGlnSerPheGlyGlnSer----- 894
QY 4413 TTTAAATTTTAAAGCAAAATCAATCAATATGAAATGAGCTTAATATATATATTA 4472
```

Thu Nov 3 08:49:23 2005

us-10-664-859-1.rapb

Page 4

we have done rfs
5482-5601
enc aa 1140-1174 ?

Db 894 ----- 894
QY 4473 TAACTGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATTAATTATTTTCCGACCG 4532
Db 894 ----- 894
QY 4533 CCCATCCATTTCTTGAAATCTTGTGAGTGTAGAGATTAATCTACTATTAAAA 4592
Db 894 ----- 894
QY 4593 TTAACACGAAAATTGATATCGTTAATGAAAATCACTATGTTTAATAGAAATTA 4652
Db 894 ----- 894
QY 4653 AATATGTTTATTATATTTCTACAGGTGATACATGAAAAGTAGCCGACCAAGCCAC 4712
Db 895 ----- -Ala--GlyAspAsnMetLysSerAlaArgProSerProG 907
QY 4713 AGGGTCAGCGGTGACAGTAATAGTCTAATAGAGCAATAAAGATGTAACGATTTGCTG 4772
Db 907 InGlyAlaArgSerProValAsnSerLeuIleGluAlaAsnLysAspValArgPheAla 927
QY 4773 CATCCAGTCTGTTTAAACCCGATCCACATATGCAAGCAATTCGAATTCAGCTTAA 4832
Db 927 LAserSerProGlyPheAsnProHlaProHlaMetGlnSerAsnSerAlaLeuA 947
QY 4833 AGCCCTATAAATGGGCTCTACCAATATACAGATGAGGTAATTTAATATTATTATT 4892
Db 947 snAlaLysLysMetGlySerThrAsnIleGlnMetGlu--Arg----- 960
QY 4893 TAAAGTTTGTGTAATTTATCTTTTTCAGCGTCAAGCATCAGCGCAAGTGATC 4952
Db 961 ----- -GlnAlaSerAlaGlnIleGlyIle 968
QY 4953 CGTACAATTTAGTCGGCGCTCCGATATATTCGGCTTAATCCCAATAGTGGCAATGGCC 5012
Db 968 rValGlnPheSerAlaArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgP 988
QY 5013 GCCACAAACAGATATACCAAACTTCGATCCATCTGTTTGGCAAAATGTGCCA 5072
Db 988 oProPheAsnLysMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerG 1008
QY 5073 ACAACTAACAGTTGGGTGTCAGATGGTATGTCAGCCGAATCGTGGTATGACAT 5132
Db 1008 ngInLeuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyIleMetThr 1028
QY 5133 GATGGGGGTCGGGACCGTCCGACATCAATATTGACATGAAATTTCCGACTTAA 5192
Db 1028 tMetGlyGlyProGlyProSerAspIleAsnIleGlnIleGlyIleIleSerGlyLeuAs 1048
QY 5193 TGGATCAGGAATGATACCAATAATCAAAATACTGATTCATCAATATGCTATATGA 5252
Db 1048 pGlySerGlyIleAspThrIleAsnGlnAsnAsnCysHlaSerMetAsnValValMetAs 1068
QY 5253 CTCATATGGGTCCCGAATGCTGAATCTTAATATGTCGTAAGCAGCGGTCCAAATGACC 5312
Db 1068 nSerMetGlyProArgMetLeuAsnProLysMetCysValAlaGlyIleProAsnGlyL 1088
QY 5313 GCGTGGCTTTAATCCCAATTCGCCCAATGGTGAATTAAGAGAAATTCATAGGCTGCG 5372
Db 1088 oProGlyPheAsnProAsnSerProAsnGlyLysLeuArgGluAsnSerIleGlySerG 1108
QY 5373 CTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGTGTGTTCCACTGTGGCAGATGAT 5432
Db 1108 yCysGlySerAlaAsnSerSerAsnPheGlnGlyValProProGlyValAlaArgMetHe 1128
QY 5433 GGGTCGAATGCCAGTCAATTTGGTTCGAATTTCAATCCGAATATTCAGGTAAGCGCAG 5492
Db 1128 tGlyArgMetProValAsnPheGlySerAsnPheAsnProAsnIle----- 1143
QY 5493 TACCCGAACACATCATATCATGTCAGTAAGGACAGAACGCCAACAAATTAACA 5552

Db 1143 ----- 1143
QY 5553 CAATGAGCTAATATATGTCGAATGCGACCTAGTCTGAATTTTTCAGAGGAGCTAA 5612
Db 1144 ----- -GlnArgTyrAlaAs 1148
QY 5613 CCTCAAAATGGGTGCTGTAAGCAATGGGTGCGCAATATGCCACATCAGCAGCGACG 5672
Db 1148 nProGlnMetGlyAlaValAlaGlyAsnGlySerProIleCysProProSerAlaSerArg 1168
QY 5673 TACTCTGGAATGCCAGATTTGATGCGGGACACAGAGCCGAGATGCTAATGAATTC 5732
Db 1168 yThrProGlyMetProGlyLysMetAlaGlyProGlyAlaGlyLysMetLeuMetAsnSe 1188
QY 5733 TTCGGAGAGCAACACCAAGATCAACAATCTCGGGGCAAGCAATGATGATTA 5792
Db 1188 tSerGlyGlnGlnIleGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAs 1208
QY 5793 CTTCTTTCAGAAATTCGAATCAATGCTATGTTGACGAAGAGGTGATTAACCGGCA 5852
Db 1208 nPhePheGlnAsnCysAsnGlnMetSerIleValAspGlnGlnGlyLysProGlyL 1228
QY 5853 TGAAGATCAATGAATATTTGTCACCAATCTATGTAAGGGGCAATCGTCCATGCCAT 5912
Db 1228 sAspGlySerMetAsnIleGlyGlnProSerMetIleAspGlyMetAspProHlaAlaMe 1248
QY 5913 GCGGCCAAATGTAAATGGGTGCGCGGATGCGACCCGTTAACAGCAAAATTCAGTTGACA 5972
Db 1248 tArgProAsnValIleMetGlyAlaArgMetProProValAsnAsnGlnIleGlnPheAlaG 1268
QY 5973 GTTCATCGATGGTATTTGACTGTGTGCGGGATCCGTGATCATTTTTCATTAACGCTTCG 6032
Db 1268 nSerSerAspGlyIleAspCysValAlaLysProSerSerPhePheThrAsnAlaSerC 1288
QY 6033 CAACGCGCTGACACACATATTTGATCAGACACAAACAGGCCAAATCAGCTTAAGCAC 6092
Db 1288 sAsnSerAlaGlyProHlaMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrG 1308
QY 6093 ACACATAAAGAACATCTAGTGAATGTGTCAAAACCAATCGGACTTGACAGTGCACA 6152
Db 1308 nHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaG 1328
QY 6153 AGCGCAGATCCAACTGATGGCGAAGACATGCGCAGGCTCAGTCTTTAATGACCTAC 6212
Db 1328 ngIyGlnIleGlnLeuIleGlyGlnIleYnIleAlaGlnIyGlnIleSerLeuIleGlyProH 1348
QY 6213 TAATAATATTTAATGTCACAGTGGGAGGTGTCAGTGTACTTAACGAGTCTGGCAT 6272
Db 1348 rAsnAsnAsnLeuMetSerThrAlaGlySerValIleSerAlaThrAsnGlyValIleSerGly 1368
QY 6273 CAATTCGTAGGTCCCTCTTCTACGACCTGAAGTATGCCAGCAATATCATATGTTTGA 6332
Db 1368 sAsnPheValGlyProSerSerThrAspLeuLysTyrAlaGlnGlnIyHlaSerPheG 1388
QY 6333 GACAGCATTAATGCTAACCAACACAGAGTCAACAAACAGCATATGCACAGACGA 6392
Db 1388 ngInGlnLeuTyrAlaThrAsnThrArgSerGlnGlnGlnIleAsnHlaGlnIleHla 1408
QY 6393 CCAAGCAACATGATTAACAATGCGCGCAATTTATACCAAAATCCAAAGTCTTTGTCA 6452
Db 1408 sGlnSerAsnMetIleThrMetProAsnLeuSerProAsnProThrPhePheValAs 1428
QY 6453 CAAA 6456
Db 1428 nLys 1429

RESULT 2

US-09-915-543-10

; Sequence 10, Application US/0915543

; Publication No. US20020086986A1

; GENERAL INFORMATION:

; APPLICANT: BASLER, Konrad

; APPLICANT: BRUNNER, Erich

```

; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Drosophila 1gs
US-10-322-579-10

Alignment Scores:
Pred. No.: 2e-38 Length: 112
Score: 570.00 Matches: 112
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.74% Indels: 0
DB: 14 Gaps: 0

US-10-664-859-1 (1-6909) x US-10-322-579-10 (1-112)

QY 4029 TCTGCGTCAGTACCAATGACCACTCAATCGCCATCCCTCGAGTCGCAACATCTATCT 4088
Db 1 SerIaSerValProlIeAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

QY 4089 CTCCCGTCACCGCGGACCAACCGCAGCACTGAGGATTCGCGACCACTTCTTACGATG 4148
Db 21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

QY 4149 GATGGAACAGATCTATTATCTCGATCTGTTCCGCAAGCTATCTACTTCACGGTTACGCA 4208
Db 41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

QY 4209 GGCAACAACAAGTGCCTCTGACAAACAAGAACTGTTTTCAGGCGACACCCCATCGCG 4268
Db 61 GlyThrThrThrValLeuSerAlaAsnLysAsnCySpneGlnAlaAspThrProSerPro 80

QY 4269 TCAATCAAAATCGTAGTAGAAATACCGATCGTCAAGCGTTCCTTACGCATTACTTAAGC 4328
Db 81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

QY 4329 AGCAACCAAGTACCCCTTATCTCATCTATCCCCA 4364
Db 101 SerAsnProSerThrProLeuSerHisLeuSerPro 112

RESULT 4
US-10-664-859-10
; Sequence 10, Application US/10664859
; Publication No. US20040038901A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMP, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/664, 859
; CURRENT FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Drosophila 1gs
US-10-664-859-10

Alignment Scores:
Pred. No.: 2e-38 Length: 112
Score: 570.00 Matches: 112
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.74% Indels: 0
DB: 14 Gaps: 0

```

DB: 15 Gaps: 0

US-10-664-859-1 (1-6909) x US-10-664-859-10 (1-112)

QY 4029 TCTGCGTCACTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTTATCT 4088
 DB 1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAlaLeuSer 20
 QY 4089 CTCCTGCTACCGCGGACCAACCGCAGAGTATGGGATGGGATGGGACCAACTCTCTAGCAAG 4148
 DB 21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40
 QY 4149 GATGGAACAGATCATTTATCTGGATCTGTTCCGCAAGCTATATCTTGAAGGTTTCAAGCA 4208
 DB 41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60
 QY 4209 GGCACACACAGAGTCTCTTACAGAACAGAACTGTTTTCAGCAGACACCCCATGCGCG 4268
 DB 61 GlyThrThrThrValLeuSerAlaAsnLysAsnCyPheGlnAlaAspThrProSerPro 80
 QY 4269 TCAAAATCAAAATGCTAGTAAATACCGGATCGTCAAGCGTTCTTACGCACTTAAAGC 4328
 DB 81 SerAsnGlnAsnArgSerAlaGlnAsnThrGlySerSerSerValLeuThrAlaAsnLeuSer 100
 QY 4329 AGCAACCAAGTACCCCTTATCTCATCTATCCCA 4364
 DB 101 SerAsnProSerThrProLeuSerThrAlaLeuSerPro 112

RESULT 5

US-10-276-774-2178
 ; Sequence 2178, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang, Y. Tom et al
 ; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21272-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; PRIOR FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: Custom
 ; SEQ ID NO 2178
 ; LENGTH: 1435
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-276-774-2178

Alignment Scores:

Pred. No.: 3,14e-17 Length: 1435
 Score: 321.50 Matches: 313
 Percent Similarity: 30.10% Conservative: 180
 Best Local Similarity: 19.11% Mismatches: 522
 Query Match: 2,67% Indels: 624
 DB: 15 Gaps: 71

US-10-664-859-1 (1-6909) x US-10-276-774-2178 (1-1435)

QY 2412 TCAAAAGCATGATCAACGAAAAAAGATCGTCTTGCATGATGATATGACGAAATG 2471
 DB 55 AsnThrGlnSerSerProLysSerLysGlnGlnValMetValArgProProThrValMet 74
 QY 2472 AGCATGAAAGCTGCAATCATGTTGAATCCCGATTTTATCATGATCTTTAAATAATCTCT 2531
 DB 75 SerProSerGlyAsnProGlnLeuAspSerLysPheSerAsnGlnGlyLysGlnGly 94
 QY 2532 GCAATTTCAGCATATTAGTAAGC----- 2555
 DB 95 SerAlaSerGlnSerGlnProSerProCyAspSerLysSerGlyGlnHisThrProLys 114

QY 2556 -----GAGTAGACCAATATACCCGGAATCGAGTGGAGCGGCGAAGAAATTTA 2606
 DB 115 AlaLeuProGlyProGlyGlySerMetGlyLeuLysAsnGlyAlaGlyLysGly 132
 QY 2607 TTGACTGCCAAGCCCAATGGA-----ATCTCTGGGCTAGC----- 2642
 DB 133 -----AlaLysGlyLysGlyLysArgGlyArgSerIleSerAlaAspSerPheArgGln 150
 QY 2643 -----AGTAATTGTTGGATTAC 2660
 DB 151 ArgAspProGlyThrProAsnAspAspSerAspIleLysGlyLysSerAlaAspHis 170
 QY 2661 ATGCAACAGCAAAATCAC----- 2678
 DB 171 IleLysSerGlnAspSerGlnHisThrProHisSerMetThrProSerAsnAlaThrAla 190
 QY 2678 ----- 2678
 DB 191 ProArgSerSerThrProProHisGlyGlnThrAlaThrGlyLysProThrProAlaGln 210
 QY 2679 -----ATATTGCGTTTTCAACTCAGCTGGCGCAACAAAGGCGCGAA 2720
 DB 211 LysThrProAlaLysValValValPheSerThrGlyMetAlaAsnLysAlaAlaGln 230
 QY 2721 TCAGTTTAAAGCGGCAATTTCAACTATTAATTCGATACCTGACCTGACCTGCTACA 2780
 DB 231 AlaValLeuLysGlyGlnValGlnThrIleValSerPheHis----- 284
 QY 2781 AAAAGCTTCTGGAGACCTTTTATGAAAAAACCTTTTAAAGATTACAGATTACAGCGG 2840
 DB 245 -----IleGlnAsn-----IleSerAsnAsnLysThrGlyArg 285
 QY 2841 CAAATTCGCGGTATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2900
 DB 256 SerThrAla-----ProLeuAsnThrGlnIleSerAlaLeuArgAsnAspProLys 292
 QY 2901 CCTGTAGCCAAATTAACACAGACGCGCATACAAAGACCGTAGGCTATTGAAACC 2960
 DB 273 ProLeu-----ProGlnGlnProProAlaProAla----- 282
 QY 2961 CAATTCAATCAATGAAAAACGCAAGCAGTAGTACTGTAAAGCGCGCTAGCAACTCTTTT 3020
 DB 283 -----AsnGlnAspGlnAsnSerSerGlnAsnThrArgLeuGlnProThrProProIle 300
 QY 3021 -----GTGACCAAGTCTGATCCTATG 3041
 DB 301 ProAlaProAlaProLysProAlaAlaProProArgProLeuAspArgGlySerPro--- 319
 QY 3042 GCGCAACGAACCTGAATTTGATGCTGGAGGCGGATCCTCAACACCACT----- 3092
 DB 320 GlyValGlnAsnLysLeuLeuProSerValGlySerProAlaSerSerThrProLeuPro 339
 QY 3093 -----AGCTCGGACAAATCTACGAAATCATGTAGACAGTATCATTCACAGCAG 3146
 DB 340 ProAspGlyThrGlyProAsnSerThrProAsnAsnArgAlaValThrProValSerGln 359
 QY 3147 TCACAGCGCAATAAGTACTGACGAGACGCTGGCGTTGATTGGGACAGCGCAAAAGGA 3206
 DB 360 GlySerAsnSerSerSerAlaAspProLysAlaProProProProProValSerSerGly 379
 QY 3207 AGCATCTGCGCTGCAACCTGAAACCAATTTGATCATCTGCAAGAGATTAGTTGATTCGA 3266
 DB 380 GluProProThrLeu---GlyGlnAsn-----Pro 388
 QY 3267 GACGAAACCTTACACACACACAGCGCAACATCGGGAAGAACAGTGGCAAAATTAATA 3326
 DB 389 Asp---GlyLeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 407
 QY 3327 AAAAGTAATCAATTTCTTTCTGTAATGGAATTCAGTAGGAGCATATGTAAGCTCA 3386
 DB 408 AspIleGlnArgMetLeuPheProAspGlyLysGlyLysThrGlyAlaGln----- 424
 QY 3387 CAGATTAACAAAAATTCACAGAGATTATGATGAGGAGTGTGGGTGGCGAGCGGATCT 3446

13	222	1.8	140	4	US-09-125-635-4	Sequence 4, Appl1
14	219	1.8	2451	4	US-09-949-016-9675	Sequence 9675, Appl1
15	216.5	1.8	1061	3	US-08-701-154A-5	Sequence 5, Appl1
16	204	1.7	1584	3	US-09-457-040B-27	Sequence 2, Appl1
17	203	1.7	2314	3	US-09-268-347-49	Sequence 49, Appl1
18	202	1.7	1600	2	US-08-617-697-10	Sequence 10, Appl1
19	201	1.7	1036	3	US-08-891-640-3	Sequence 3, Appl1
20	201	1.7	1302	4	US-09-949-016-110852	Sequence 10852, Appl1
21	198.5	1.6	1258	4	US-09-538-092-886	Sequence 886, Appl1
22	198.5	1.6	1531	4	US-09-949-016-6740	Sequence 6740, Appl1
23	198	1.6	2703	1	US-08-188-432-19	Sequence 19, Appl1
24	198	1.6	2703	4	US-08-899-232-4	Sequence 4, Appl1
25	198	1.6	2703	4	US-09-121-457-4	Sequence 4, Appl1
26	196.5	1.6	2504	4	US-09-328-352-5821	Sequence 5821, Appl1
27	194.5	1.6	1529	2	US-08-728-470-10	Sequence 10, Appl1
28	194.5	1.6	1529	2	US-08-719-641-10	Sequence 10, Appl1
29	193	1.6	908	4	US-09-949-016-6542	Sequence 6542, Appl1
30	193	1.6	2189	4	US-09-949-016-6930	Sequence 6930, Appl1
31	192	1.6	2736	4	US-09-252-991A-30227	Sequence 30227, Appl1
32	191	1.6	3969	3	US-08-061-376-5	Sequence 5, Appl1
33	191	1.6	3969	4	US-09-538-092-1262	Sequence 1262, Appl1
34	189	1.6	1268	4	US-09-270-767-15320	Sequence 15320, Appl1
35	186.5	1.5	1776	4	US-09-556-877-179	Sequence 179, Appl1
36	186.5	1.5	1776	4	US-09-620-412C-179	Sequence 179, Appl1
37	186.5	1.5	1776	4	US-09-598-419-179	Sequence 179, Appl1
38	183.5	1.5	1187	4	US-09-949-016-6513	Sequence 6513, Appl1
39	183.5	1.5	1402	4	US-09-448-353E-2	Sequence 2, Appl1
40	181.5	1.5	1477	4	US-09-206-942-71	Sequence 71, Appl1
41	180.5	1.5	1992	4	US-08-621-944A-3	Sequence 3, Appl1
42	180.5	1.5	1992	4	US-08-945-567D-3	Sequence 3, Appl1
43	179.5	1.5	1833	4	US-08-621-944A-4	Sequence 4, Appl1
44	179.5	1.5	1833	4	US-08-945-567D-4	Sequence 4, Appl1
45	176.5	1.5	1833	4	US-09-248-567D-4	Sequence 4, Appl1
			1232	4	US-09-949-796A-19543	Sequence 19543, Appl1

ALIGNMENTS

US-09-270-767-44797 ✓
 : Sequence 44797, Application US/09270767
 : Patent No. 6703491
 : GENERAL INFORMATION:
 : APPLICANT: Homuburger et al.
 : TITLE OF INVENTION: Nucleic acids and proteins
 : FILE REFERENCE: File Reference: 7326-094
 : CURRENT APPLICATION NUMBER: US/09/270,767
 : CURRENT FILING DATE: 1999-03-17
 : NUMBER OF SEQ ID NOS: 62517
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 44797
 : LENGTH: 506
 : TYPE: PRT
 : ORGANISM: Drosophila melanogaster
 : FEATURE:
 : OTHER INFORMATION: Xaa means any amino acid.
 : US-09-270-767-44797

Programatic Occurs:	
Pred. No.:	6.03e-220
Score:	2459.50
Percent Similarity:	80.03%
Best Local Similarity:	79.71%
Query Match:	20.18%
DB:	4
Length:	506
Matches:	499
Conservative:	2
Mismatches:	9
Indels:	116
Gaps:	2

US-10-664-859-1 (1-6909) x US-09-270-767-44797 (1-506

3271 - AAAACCTTACACCAACAGCGCAACATCGGGAAGAACAGTTGGCAAAATAAAAAAA 3330

Db 1 LysThrLeuHisHisAsnSerGlyAsnIleGlyLysAsnSerTrpGlnLys**LysLys 20

3331 TGAATCAATTTCTTTTCTGAAATGAGATTTCAGTAGGAGCTAATGTAAGCTCACAGA 3390

```

Db      21 ***IleAsnPhenHeuHeuMetArgIleGln**GluLeuMet**AlaHisArg 40
QY      3391 TAAACAAAATTCAGAGATTTATATGATGGAGATGCGGGTGGCAGAGCGGATCTATT 3449
Db      41 ***GlnIlePheGlnIleIle*****TyrGlyCysGlnValAlaGlyGlyGlySerIle 60
QY      3450 ATTAATCCAGACGATGCGACAACTGATATGCGAGGTAAACGCCAAATCGAGGCTTTATGC 3509
Db      61 IleAsnProThrMetArgIleuHisMetProGlyAsnAlaIleSerGluLeuLeuSer 80
QY      3510 GCGACAGTTCCAGACCTTCCGAGAGATGTAATCATCCAGGGGATGTTATATCAGATAG 3569
Db      81 AlaThrSerSerGluLeuSerGluAspValMetHisProGlyAspValIleSerAspMet 100
QY      3570 GGTGCGGTATATAGATGTAATTAATTAATCAAAAAACAGTGTGCAATGTGATCTGAGTA 3629
Db      101 GlyAlaValIleGlyCysAsnAsnAsnGlnIleThrSerValGlnCysGlySerGlyVal 120
QY      3630 GGTGTGTCTACCTGAAACAACTGCACTGAGTAAATGTCATATGCAATGCTCAAGCTTC 3689
Db      121 GlyValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSer 140
QY      3690 GGGGCCCCGGAATGGCAATATATGATGGAGAGCTTACGATATGCTAGCCTCGTTGGCAAC 3749
Db      141 GlyAlaProAsnGlnIleMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsn 160
QY      3750 ACAAGTCGCAACGTCATCGGACCGGCCAGATATGCTTAAGAGATTTTAATCAAGAT 3809
Db      161 ThrSerCysAsnValIleGlyThrAlaProAspMetSerIleGlyValIleuAsnGlnIle 180
QY      3810 AGCGGAACCATTCACATCAAGGGGAGTGGTAAATGAGTGGTCAAGATCAACAT 3869
Db      181 SerArgThrIleSerHisGlnGlnGlyValAlaGlnMetCylurPserIleGlnIleHis 200
QY      3870 CAATTTTGAAGAAGCTTCAAGGGGGGAGCCGAGCAAGCACTGGAACCTGTAAGTA 3929
Db      201 GlnPhePheGlnGlnIleGlnGlyGlyValProGlnGlnValIleThrValIleVal 220
QY      3930 CCACAAACAGCAAAACCCCTTCTGATCTGTGTAAGTCTGTTAAACACAGTGCAGACC 3989
Db      221 ProGlnGlnIleThrProSerGlySerGlyValAsnSerIleuAsnGlnIleValArgPro 240
QY      3990 CTGCAAGTCCACCTCTCTCTTACCATTCATCAAGATCTGCGTCAAGTCAACATAGCC 4049
Db      241 LeuGlnGlyProProProProGlyIleSerIleGlnAlaSerIleAsnValProIleAla 260
QY      4050 ACTCAATCGCCCAATCCCTGAGTCCCAACCAATCTATCTCTCCGTCACCGGAGCAAC 4109
Db      261 ThrGlnSerProAsnProSerSerProAsnAsnLeuSerIleuProSerProAlaGlnThr 280
QY      4110 GCAGCACTCATGGATTCGCAACCAACTCTCTAGCATGATGAAACAGATCATTAATCT 4169
Db      281 AlaAlaValMetGlyLeuProThrAsnSerProSerMetHisPheGlyThrGlySerLeuSer 300
QY      4170 GGAATCTGTCCGACGATATCTTGAACGCTTCAAGGACGACACACAAACAGTCTCTCA 4229
Db      301 GlySerValProGlnAlaAsnThrThrValGlnAlaGlyThrThrThrValLeuSer 320
QY      4230 GCAAAACAAGAACTGTTTTCAGGACAGACACCCCAATCGCGTCAATCAAAATCGTAGAG 4289
Db      321 AlaAsnIleAsnGlyPheGlnAlaAspThrProSerProSerAsnGlnAlaAspSerArg 340
QY      4290 AATACCGGATGTCACGCTTCTTACGATATCAATTAAGCAACCAACCAATGACCCCTTA 4349
Db      341 AsnThrIleSerSerSerValIleuThrHisAsnLeuSerSerAsnProSerThrProLeu 360
QY      4350 TTCATCTTATCCCAAAAGAAATTTGAGTCTTTGCGTAGTCTCTGCTGATGTTATATAT 4409
Db      361 SerHisLeuSerProGlyGlnPheGlnSerPheGlyGlnSerSer 375
QY      4410 TTGTTAATTTTAAAGCAAAATCAAAATGAAATGCGTTAATATATATATATATATATAT 4469
Db      375 ----- 375

```

```

QY      4470 ACATAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATTAATTTTCCGGA 4529
Db      375 ----- 375
QY      4530 CCGCCCATTCATTTCTTGAATCCAAATTTCTGAGATGATTTGATAGATTAATCTACTATTA 4589
Db      375 ----- 375
QY      4590 AAATTAACACGAAATTCATATCCGTTAATTAATAATCACTATGTTTAATAGAAATTT 4649
Db      375 ----- 375
QY      4650 AAAAAATGTTTATTAATTAATTTCTACAGGTATTAACATGAAAAAGTACGACCAAGCC 4709
Db      376 ----- 376
QY      4710 CACAGGTCAGGGTCCACCACTTAATTAATGATGAGCAAAATTAAGATGATGATTTG 4769
Db      387 TGGGlnGlyGlnArgSerProValAsnSerLeuIleGlnAlaAsnIleAspValArgPheAla 407
QY      4770 CTGATCCAGTCTGCTTTTAAACCGCATCCCATATGCAAAAGCAATTCATATGAGCAT 4829
Db      407 AlaIleSerSerProGlyPheAsnProHisProHisMetGlnSerSerAsnSerAlaAla 427
QY      4830 TAAAGCCCTATTAATTAATGGGCTCTTACCAATATACAGATGAGGTAAATTAATATATTT 4889
Db      427 euAsnAlaIleIleMetGlySerThrAsnIleGlnMetGly-Arg 441
QY      4890 ATTTAAGCTTTTGTGTTAATTAATTAATCTCTTTTACGCTCAAGATGACGCGCAAGTGC 4949
Db      442 ----- 442
QY      4950 ATCCGTACATTTAATGCGCGCTCCGATATATATTCGCTTAATCCCAATAGTGCAGATCG 5009
Db      448 ySerValGlnPheSerArgArgSerSerAspAsnIleProIleuAsnProAsnSerGlyAsnArg 468
QY      5010 GCGGCCACCAACAAAGATGACCCCAAACTTGATCCATCTCTTTGGCAAAATGTC 5069
Db      468 GlnGlnIleuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGly 506
QY      5070 CCAACACATTAACAAATTTGGGTCTCCAGATGAGTGTGTCAGCGGCAACGTGGT 5124
Db      488 TglnGlnIleuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGly 506

RESULT 2
US-09-270-767-60260
; Sequence 60260, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 60260
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60260

Alignment Scores:
Pred. No.: 2.8e-139
Score: 1573.00
Percent Similarity: 97.48%
Best Local Similarity: 97.16%
Query Match: 13.07%
DB: 4
Length: 317
Matches: 308
Conservative: 1
Mismatch: 8
Indels: 1
Gaps: 0

```


US-10-664-859.1 (1-6909) x US-09-270-767-60260 (1-317)

QY	3271	AAACCTTACACACACAGCGGCAACTCGGAGAAACAGTTGGCAAAATAAAAA	3330
Db	1	LySTrLrLwHnShnIaSnSerClYAnIlIeGlYbAnSerTrpGlInLys**LysLys	20
QY	3331	TGAATCAATTCCTTTCTGTAATAATGAGATTCTGAGAGCTTAATGATGCACGA	3390
Db	21	**IleaSnPhePhePheLeuLysMetArgIleGln**GluLeuMet**AlaIbArg	40
QY	3391	TAAACAAAATTCACAGAGATTTAATGATGGGATTCGGGTGC-CGAGCGGATCTATT	3449
Db	41	***GlnLysPheGlnGlnIle*****TrpGlyCysGlnValAlaGlyIleGlySerIle	60
QY	3450	ATAAATCCGACGATCGCAACATGTCATATGCCAGGTAAAGCCAAATCGACCTTTATGC	3509
Db	61	IleaSnProthnMetArgGlnLeuHisMetCProGlyAlaValAlaLysSerCluLeuLys	80
QY	3510	GGCAACATTCAGGACTTTTGGAAATGTAATGCAATCCAGGGAGTGTATATCAGATATG	3565
Db	81	AlaTrnSerSerGlyLeuSerCluAspValMetHisProGlyAspValIleSerAspMet	100
QY	3570	GGTGGCGTAATAGATGTAT	3629
Db	101	GlyAlaValIleIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerCluVal	120
QY	3630	GATGTGTCACTGGAACAATCGACCTGAGATTAATGTCAATATGTCATATGCTCAAGCTCC	3689
Db	121	GlyValValIlnGlnGlyThrThrAlaIleGlyAlaValAlaValAlaMetHisCysSerSer	140
QY	3690	GGCGCGCCGGAATGGCAATATGATGGGAAGCTCTACGATATGCTATGCTGTTGGCAAC	3749
Db	141	GlyAlaProAsnGlyAsnMetMetClySerSerThrAspMetLeuAlaSerPheCluAsn	160
QY	3750	ACAAGCTCAACGTCATCGAAGCGGCCACGATATGTCATAGAAAGTTTAAATCAACAT	3809
Db	161	ThrSerCysAsnValIleIleGlyThrAlaProAspMetSerLysGlyValLeuAsnGlnAsp	180
QY	3810	AGCGGAACCCATTCACATCAAGAGGGAGTTGCTCAATGAGATGGTCCAGAGATTCACAT	3869
Db	181	SerArgThrHisSerHisGlnGlyLysIleValAlaIleGlnMetGlnTrpSerLysIleGlnHis	200
QY	3870	CAATTTTTCGAAGAACGCTTCAAGGGGGGCAAGCCACAGACATGTCATGTAATCTGTAGTA	3929
Db	201	GlnPhePheGlnGlnArgLeuLysGlyLysProArgGlnValIleThrGlyThrValVal	220
QY	3930	CCACAAACGCAAAACCCCTTCTGGAATCTGAGTGGAAACCTGTTAAACAACACAGATCGAACCC	3989
Db	221	ProGlnGlnGlnIlnHisProSerGlySerIleGlyLysAsnSerLeuAsnAsnGlnValArgPro	240
QY	3990	CTGCAAGGTCAACCTCTCTTACACATCCATCCAGAGATGCGTCAAGTACCAATAGCC	4049
Db	241	LeuGlnGlyProProProProGlyHisSerIleGlnArgSerAlaSerValProIleAla	260
QY	4050	ACTCAATGCCCAATCCTCGAGTCCAAACAATCTATTTCTCCCGTCAACCGGCAACAC	4109
Db	261	ThrGlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThr	280
QY	4110	GCAGAGTCATGGGATTTGCGACCAACCTCCTTGACATGATGGAATGGAAACGATCATTTATCT	4169
Db	281	AlaAlaValMetGlyLeuProThrAsnSerProSerMetAspGlyIlnGlySerLeuSer	300
QY	4170	GGATTCGTTCGCAAGCTAATATCTGACGGTTCAGGACGACGACACAAACA 4220	
Db	301	GlySerValProGlnAlaAsnThrSerThrValGlnAlaGlyIlnThrThr 317	

RESULT 3
US-09-270-767-46113
; Sequence 46113, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*

```

; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46113
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46113

Alignment Scores:
Pred. No.:      5,836-99      Length:      259
Score:          1145.00      Matches:      250
Percent Similarity: 50.60%      Conservative: 1
Best Local Similarity: 50.40%      Mismatches: 1
Query Match:      9.51%      Indels:      244
DB:                4      Gaps:      1

US-10-664-859-1 (1-6909) x US-09-270-767-46113 (1-259)

QY      1453 TCAGAAATTAAGTTCTAATAAAGCAAAAGTCAACACCTGGTCGCGCTCGCAACAGCT 1512
Db      8 Aaagluileseersemuysalalygilyglnalaaaglygilycyvgluhrclly 27
QY      1513 TCTACATCCAGTTTGACTGCTCAAGAGAACCCACGATGCTTAGAGCAGTTTAGTAAT 1572
Db      28 SerthrseerleuthrivalysglugluProthrspallLeuglyserLeuValasn 47
QY      1573 ATGAAAAAAGAGAAAGAGAAATCATTTCCGCAACGATGTCGCCCTGTGGTTTGTGTTCA 1632
Db      48 MetylslysglugluAagluasnHiserProthrnetserProvalGlyPhleglyser 67
QY      1633 ATTGGTAATGCACGACGACAACTCCGCTACACCGGGTAAGTTTAAAGATTCATATAAG 1692
Db      68 lileglyasnlaaglnaepasnseralathr----- 77
QY      1693 CAATAACAGAAATTAATGTCAGTTACCAATTTTATTTGATAGTCAAAAGACTACTATAG 1752
Db      77 ----- 77
QY      1753 CGATATCTCCTGCCTTTTAAATTTTATTTAATAGAAATACGAATATTTCTAATTTGTA 1812
Db      77 ----- 77
QY      1813 AAATAAAATTGATTAAATTAAGTAATTTAAAAACCTTTGAATTAGACATACCCTTCC 1872
Db      77 ----- 77
QY      1873 AAAAAATCAGTAATCATTTGGGAACGAGAGTGTGTCGCCGAAGAGACTACTATAAACCTT 1932
Db      77 ----- 77
QY      1933 TTGAGCTATCTGATTAAGTCAACGCTACTAAAAAATGATTAGTTAGAAAAATGGGTAAAT 1992
Db      77 ----- 77
QY      1993 TTGTAGGAAGTTTCAATTTTGAAGAAGATGTGATTATTTATTAACCCCTTCAAGCGGA 2052
Db      77 ----- 77
QY      2053 ACTACATTTGTTCTACGATATTTTGGAAAAACAATGTTAAGTTGGAAGTGCCTATAA 2112
Db      77 ----- 77
QY      2113 AACGAATTCACGCGTTTCAAAATCTAACACAGGTTTGTATTAAATTTGATTAAATGAG 2172
Db      77 ----- 77
QY      2173 AAATTATCACACTTCAATTAAATGTTTAATTTGATTAAGGTGGACAATCACAGAGAT 2232
Db      77 ----- 77

```

```
QY 2233 TTCATTTTTCGCTGATATATATAGAAAGTCGCTTCACTCTTTCGCGCGCTTCAACAC 2292
Db 77 ----- 77
QY 2293 TACGTGAGATTCGCCCGACGATTTATATAGATGATTTAGAGATTATTTTATTTT 2352
Db 77 ----- 77
QY 2353 TGGTGATTTTAAATAATCTTATTATTCATTTTACATAGTTAAATTTGAAAGATTT 2412
Db 78 -----Pro--ValLysIleGluArgIleS 85
QY 2413 CAACGACAGTACACGAGAAAAAAGATCGTCTTGACATAGATTAATGACGAATGCA 2472
Db 85 eraAnaSpSerThrTrhGlnLysAlaSerLysSerSerLeuThrMetAenAnaSpGluMetS 105
QY 2473 GCATGGAAGGCTGCATCAGTTGAATCCGATTTTTCATGATGATCTTTTAAATATCTGT 2532
Db 105 erMetGluGlyCySaenGlnLeuAnaSpProAspPheIleAnaGlnSerLeuAnaSpProA 125
QY 2533 CAATTCGACATATATTAGTAAGCGAGTAGACCAATACCCGGAATCGAGTTGAGCG 2592
Db 125 laIleSerSerIleLeuValSerGlyValGlyProIleProGlyIleGlyValIleG 145
QY 2593 GGAACGGAAATTTATTTGACTGCCAAGCCCAATGGAATCTCTCGGAGTACGATTTGTT 2652
Db 145 lYthrGlyAnaLeuLeuThrAlaAnaAlaAnaGlyIleSerSerGlySerSerAnaCybe 165
QY 2653 TGATTTACATGACAAACAGCAAAATCACATATTCGTTTTCATCTCAGCTGCGCAACAA 2712
Db 165 euAnaPyrMetGlnGlnGlnAnaSpHlePheValIlePheSerTrhGlnLeuAlaAnaLys 185
QY 2713 GGGCCGAATCAGTTTAAAGCGATCAATTTCAATTTATTTGCTATTCATCTGACCTCAG 2772
Db 185 lYlaGlnSerValIleuSerGlyGlnPheGlnTrhIleAlaLysCyThrGlnP 205
QY 2773 CTGCTACAAAAGCTTCTTGGAAGACTTTTATGAAAAACCTTTAAAGATTAAACAGT 2832
Db 205 roAlaThrLysSerPheLeuGluAspPheMetLysAnaProLeuLysIleAnaLysL 225
QY 2833 TACAGCGGACAAATTCGTCGATGTCATGATGATGAGGAGGAGTGAATCACTAC 2892
Db 225 euGlnTrhLysAnaSerValGlyMetProTrpIleGlyMetGlyGlnValGlyLeuThrP 245
QY 2893 CTCTATCTCTGATGCCAAATAACACACAGCAGCCACATACA 2936
Db 245 roProAnaProValAlaLysIleThrGlnGlnGlnProHisThr 259
RESULT 4
US-09-270-767-36663
; Sequence 36663, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36663
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-36663
Alignment Scores:
Pred. No.: 2,97e-79 Length: 175
Score: 935.00 Matches: 175
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.77% Indels: 0
DB: 4 Gaps: 0
```

```
US-10-664-859-1 (1-6909) x US-09-270-767-36663 (1-175)
QY 5725 ATGAATTTCTTCGAGAGACCAACCAAGATCACAAACATCTTGCGCAACAT 5784
Db 1 MetAnaSerSerGlyGluGlnIleGlnAnaLysIleThrAnaAnaProGlyAlaSerAna 20
QY 5785 GGTATTAATCTTTTCAGAAATTCATCAATATGTCATTTGTCAGGAAGGCTGATTA 5844
Db 21 GlyIleAnaPhePheGlnAnaCysAnaGlnMetSerIleValAspGluGluGlyLeu 40
QY 5845 CCCGCGCATGACGATCAATGATATGTCATCAACATCATATGATTAAGGGCATGGCTCA 5904
Db 41 ProGlyAnaSpGlySerMetAnaIleGlyGlnProSerMetIleArgGlyMetArgPro 60
QY 5905 CATGCCATGCGCCCAATGTAATGGGTGCGCGGATGCCAACCCTTAACAGCAAAATTCAG 5964
Db 61 HisAlaMetArgProAnaValMetGlyAlaArgMetProProValAnaAnaGlnIleGln 80
QY 5965 TTTGCACAGTCATCCGATGCTATTGACTGTGTCGGGAGATCCGTCATCTTTTCACTAC 6024
Db 81 PheAlaGlnSerSerAspGlyIleAspCysValGlyAspProSerSerPheThrAna 100
QY 6025 GCTTCCTGACAGAGGCTGACACACATGTTTGGATCGACACACAGGCCAATCAGCCT 6084
Db 101 AlaSerCysAnaSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAnaGlnPro 120
QY 6085 AAGACACACACATTAAGAACAATACCTAGTGAATGTGTCAAAACCAATCGGACTTGA 6144
Db 121 LysTrhGlnHisIleLysAnaIleProSerGlyMetCysGlnAnaGlnSerGlyLeuAla 140
QY 6145 GTGGCAACAGGCGAGATCCAACTGCATGAGGACAGACATGCCAGGCTGATTTAAT 6204
Db 141 ValAlaGlnGlyGlnIleGlnLeuHisGlyGlnGlyAnaGlnGlyGlnSerLeuIle 160
QY 6205 GGACCTACTAATTAATTAATTAATGTCATGCGCGAAGTGAGT 6249
Db 161 GlyProThrAnaAnaSpLeuMetSerTrhAlaGlySerValSer 175
RESULT 5
US-09-270-767-51880
; Sequence 51880, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51880
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51880
Alignment Scores:
Pred. No.: 2,97e-79 Length: 175
Score: 935.00 Matches: 175
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.77% Indels: 0
DB: 4 Gaps: 0
US-10-664-859-1 (1-6909) x US-09-270-767-51880 (1-175)
QY 5725 ATGAATTTCTTCGAGAGACCAACCAAGATCACAAACATCTTGCGCAACAT 5784
Db 1 MetAnaSerSerGlyGluGlnIleGlnAnaLysIleThrAnaAnaProGlyAlaSerAna 20
QY 5785 GGTATTAATCTTTTCAGAAATTCATCAATATGTCATTTGTCAGGAAGGCTGATTA 5844
Db 21 GlyIleAnaPhePheGlnAnaCysAnaGlnMetSerIleValAspGluGluGlyLeu 40
QY 5845 CCCGCGCATGACGATCAATGATATGTCATCAACATCATATGATTAAGGGCATGGCTCA 5904
Db 41 ProGlyAnaSpGlySerMetAnaIleGlyGlnProSerMetIleArgGlyMetArgPro 60
QY 5905 CATGCCATGCGCCCAATGTAATGGGTGCGCGGATGCCAACCCTTAACAGCAAAATTCAG 5964
Db 61 HisAlaMetArgProAnaValMetGlyAlaArgMetProProValAnaAnaGlnIleGln 80
QY 5965 TTTGCACAGTCATCCGATGCTATTGACTGTGTCGGGAGATCCGTCATCTTTTCACTAC 6024
Db 81 PheAlaGlnSerSerAspGlyIleAspCysValGlyAspProSerSerPheThrAna 100
QY 6025 GCTTCCTGACAGAGGCTGACACACATGTTTGGATCGACACACAGGCCAATCAGCCT 6084
Db 101 AlaSerCysAnaSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAnaGlnPro 120
QY 6085 AAGACACACACATTAAGAACAATACCTAGTGAATGTGTCAAAACCAATCGGACTTGA 6144
Db 121 LysTrhGlnHisIleLysAnaIleProSerGlyMetCysGlnAnaGlnSerGlyLeuAla 140
QY 6145 GTGGCAACAGGCGAGATCCAACTGCATGAGGACAGACATGCCAGGCTGATTTAAT 6204
Db 141 ValAlaGlnGlyGlnIleGlnLeuHisGlyGlnGlyAnaGlnGlyGlnSerLeuIle 160
QY 6205 GGACCTACTAATTAATTAATTAATGTCATGCGCGAAGTGAGT 6249
Db 161 GlyProThrAnaAnaSpLeuMetSerTrhAlaGlySerValSer 175
```

Db 21 GlyIleasnRhepHeGInsnCysAsnGlnMetSerIleValAspGlnGlnGlyLeu 40
 QY 5845 CCGGGCCATGAGGATCAATGATATTGTCAACCACTCTGATGTAAGGGGCGGCGCCCA 5904
 Db 41 ProGlyHisAspGlySerMetAsnIleGlyGlnProSerMetIleAspGlyMetArgPro 60
 QY 5905 CATGCCATGGCGCAATATGTAATGGTGGCGGATGCGCAACCGCTTAACAGCAAAATTCAG 5964
 Db 61 HisAlaMetArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGln 80
 QY 5965 TTTCGACAGTCATCGGATGATGTTGACTGTGTGGGGGATCCGTCATCATTTTCTACTAAC 6024
 Db 81 PheAlaGlnSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsn 100
 QY 6025 GCTTCGTCGAACAGCGGTGGACACACATGTTTGGATCAGACACACAGCGGCATTCAGCCT 6084
 Db 101 AlaSerCysAsnSerAlaGlyProHisMetPheGlySerIleGlnGlnAlaAsnGlnPro 120
 QY 6085 AAGACACACACATTAAGAACATATCACTGATGTAATGTCATAAACCATCGGACTTCGA 6144
 Db 121 LysThrGlnHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAla 140
 QY 6145 GTGGCAACAGGCAAGATCCAACTGCAATGGGCAAGACATCGCAGGGTCACTCTTAATT 6204
 Db 141 ValAlaGlnGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlnGlnSerLeuIle 160
 QY 6205 GGAACCTAATTAATTAATTAATGTCACACTGCGGGAAGTCTCACT 6249
 Db 161 GlyProThrAsnAsnMetSerThrAlaGlySerValSer 175
 RESULT 6
 US-09-270-767-44999
 ; Sequence 44999, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 44999
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 US-09-270-767-44999
 Alignment Scores:
 Pred. No.: 1,54e-47 Length: 230
 Score: 599.50 Matches: 118
 Percent Similarity: 93.13% Conservative: 4
 Best Local Similarity: 90.08% Mismatches: 6
 Query Match: 4.98% Indels: 3
 DB: 4 Gaps: 1
 US-10-664-859-1 (1-6909) x US-09-270-767-44999 (1-230)
 QY 2382 TCATTTTACATAGTAAATTTGAAGAATTTCAAGACAGTACACAGAAATAAAGAA 2441
 Db 103 SerPheIyr***SerThrValLysLeuMetSer***AsnGly-----LysLysGly 119
 QY 2442 TCGTCCTTGAACAATGAATGAAGAAATAGACATGAGAGGCTGCAATCAGTTGAATCCC 2501
 Db 120 SerSerLeuThrMetArgAsnAsnAspGlnMetSerMetGlnGlyCysAsnGlnLeuAsnPro 139
 QY 2502 GATTTTATCAATGAATTTTAAATATCTTCGCAATTTTCGACATATTAAGTAAGCGAGTA 2561
 Db 140 AspPheIleAsnGlnSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyVal 159
 QY 2562 GGACCAATACCGGAATCGAGTTTGAGCGGAGCGGGAATTAATTGACTGCCAACGCC 2621

Db 160 GlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAla 179
 QY 2622 AATGAAATCTCTCGGGGTAGCAGTAATTTGTTGATTAATGCAACAGCAAAATCACTA 2681
 Db 180 AsnGlyIleSerSerIleSerSerIleSerSerIleSerSerIleSerSerIleSerSerIle 199
 QY 2682 TTTCGTTTTCACATGAGTGGCGCAACAAAGGGGCGGAATCACTTTTAAGCGGTCAATT 2741
 Db 200 PheValPheSerThrGlnLeuAlaAsnLysGlyAlaGlnSerValLeuSerGlyGlnPhe 219
 QY 2742 CAACATTAATTAATTCGATATCACTGACCTCAGCCT 2774
 Db 220 GlnThrIleIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 230
 RESULT 7
 US-09-270-767-36764
 ; Sequence 36764, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 36764
 ; LENGTH: 140
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 US-09-270-767-36764
 Alignment Scores:
 Pred. No.: 3.8e-47 Length: 140
 Score: 594.00 Matches: 123
 Percent Similarity: 88.57% Conservative: 1
 Best Local Similarity: 87.86% Mismatches: 16
 Query Match: 4.93% Indels: 0
 DB: 4 Gaps: 0
 US-10-664-859-1 (1-6909) x US-09-270-767-36764 (1-140)
 QY 6391 CACGAGCAACATGATTAACATATGCGCGCAATTTATCAACCAATCCACGTTCTTGTG 6450
 Db 1 HisGlnSerAsnMetIleThrMetProSerAsnLeuSerProAsnProThrPheVal 20
 QY 6451 AACAAATAAATCTTAAATTTTGGCGCCCTGTCATGTAATGTTTACTAGTCTCCAAAT 6510
 Db 21 AsnLys***ThrSerLysPheLeuProSerLeuProCysIleValIyr***SerProAsn 40
 QY 6511 TAAGACATGCATCTCTAAATTAAGATTTTGAAGCTTAATTTACTAGTGTGTTTACAAC 6570
 Db 41 ***AspMetHisLeu***IleArgPhePheGlnAlaIleLeuLeuArgCysPheIyrAsn 60
 QY 6571 GGAAGAAATTAATCTTTGATATGCAATGATTAAGCTTGAAGAAACAATTAATTCATTGC 6630
 Db 61 GlyLysAsnLysLeuLeuAspMetGlnMetIleThrLeuGlnIleThr***PheIleCys 80
 QY 6631 AACTTTAGACAGTCAGTGAAGTTAAATGTAAGTAAATCTGATTTTAAACATTAAGTCAT 6690
 Db 81 AsnPhe***LysSerIleArgIser***Met***AsnLeuIyrPheAsnIleIleGlyHis 100
 QY 6691 CTGTAAATAATTAATTAACATCGAAATTTTATGATTAAGACAGCTATTTCTGTTATTAATT 6750
 Db 101 LeuLeuLys***LeuAsnIleGlnIleLeuValIleSerSerIyrIleLeuLeuPhe 120
 QY 6751 AATATGCGGCTGCTCTCTGTTGTTAAATGAATTAATTAATTAATTAATTAATTAATTA 6810
 Db 121 AsnMetCysAlaIleLeuSerValLeuAsnGlnIleIleLysIleIle***Met***Asn 140

```

RESULT 8
US-09-270-767-51981
; Sequence 51981, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 51981
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-51981

Alignment Scores:
Pred. No.: 3.8e-47 Length: 140
Score: 594.00 Matches: 123
Percent Similarity: 88.57% Conservative: 1
Best Local Similarity: 87.86% Mismatches: 16
Query Match: 4.93% Indels: 0
DB: Gaps: 0

US-10-664-859-1 (1-6909) x US-09-270-767-51981 (1-140)

QY 6391 CACCAGACGACATGATTAACAATGCGCCGAATTTATCACCAGTCTGTTGTC 6450
DB 1 HIGLISerAsnMetIleThrMetProSerAsnLeuSerProAsnProThrPheVal 20
QY 6451 AACCAATTAACCTCTAATTTTGGCCGCTCCGACATGTTGTTACTAGTCCCAAT 6510
DB 21 AsnLys***ThrSerLysPheLeuProSerLeuProCysIleValIyr***SerProAsn 40
QY 6511 TAAGACATGATCTCTAATAATAGATTTTGAAGCTTATTACTAGTGTGTTTCAAC 6570
DB 41 ***AspMetHisLeu***IleArgPhePheGluAlaTyrLeuLeuArgCysPheTyrAsn 60
QY 6571 GAGGAAATTAACCTTTGGATATGCAATATGATACGTTGGAACACATATTCATTGTC 6630
DB 61 GtGluAsnLysLeuLeuAspMetGlnMetIleThrLeuGluThr***PheIleCys 80
QY 6631 AACCTTAGAAGTCAGTCGAGTTAAGTAACTGATTTTAAACATATAGGTCAT 6690
DB 81 AsnPhe***LysSerArgArgSer***Met***AsnLeuTyrPheAsnIleIleGlyHis 100
QY 6691 CTGTAAATTAATTAACATCGAATTTAGTATGACGACTAATTTCTGTTATATTT 6750
DB 101 LeuLeuLys***LeuAsnIleGluIleLeuValIleSerSerTyrTyrLeuLeuPhe 120
QY 6751 AATATGCGCGCTGCTCTCTCTGTTTAAATGAATTAATATATATATTAAGTAAAC 6810
DB 121 AsnMetCysAlaIleAlaLeuSerValLeuAsnGluIleLysIleTyrIle***Met***Asn 140

RESULT 9
US-09-270-767-43686
; Sequence 43686, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43686
; LENGTH: 65
; TYPE: PRT

```

```

; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43686

Alignment Scores:
Pred. No.: 9.59e-21 Length: 65
Score: 312.00 Matches: 63
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 2.58% Indels: 1
DB: Gaps: 0

US-10-664-859-1 (1-6909) x US-09-270-767-43686 (1-65)

QY 4766 ATGTGACATCTTTATTTGGCTTCTATTAAGATTTACTGTGAGCCGCTGACCTTGACC 4707
DB 1 IleValHisLeuTyrLeuProLeuLeuAspTyrLeuLeuValThrAlaAspProValGly 20
QY 4706 TTGATCGCTACTTTTTCATGTTATCACTGTGAAATATTAATTAACATATTTTAAAT 4647
DB 21 LeuValAlaTyrPheSerCysTyrHisLeu***LysTyrTyrAsnLysHisIlePheAsn 40
QY 4646 TTCTATTAAACAATAGATGATTTTCATTTAAC-GGATATGAATTTTGGTTAATTTTA 4588
DB 41 PheLeuLeuAsnAsnSerAspPheGlnLeuThrGlyTyrGluPheSerCysLeuIleLeu 60
QY 4587 ATAGTAGATTATCTC 4573
DB 61 IleValAspTyrLeu 65

RESULT 10
US-09-949-016-7088
; Sequence 7088, Application US/0949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMOPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7088
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7088

Alignment Scores:
Pred. No.: 4.52e-13 Length: 1179
Score: 238.50 Matches: 250
Percent Similarity: 30.38% Conservative: 137
Best Local Similarity: 19.62% Mismatches: 406
Query Match: 1.98% Indels: 461
DB: Gaps: 64

US-10-664-859-1 (1-6909) x US-09-949-016-7088 (1-1179)

QY 2814 CCTTAAAGATTAAACAAGTTACAGCGGACCAATTCGCTGATGCGATGATAGGCATG 2873
DB 156 ProSerGlnGlnSerThrThrGlnGlnHisProGlnSerGlnProProTyr----- 172
QY 2874 GGGCAGGTGACTACTCTCTCTATCTCTGATAGCCAAATTAACAACAACAGCAGCAT 2933
DB 173 -----SerGlnProGlnAlaGlnSerProTyrGlnGlnGlnGlnProGln 187

```

OY		2994	ACAAAGCCGTAAGCGCTATTG-----MAACCSCCAATTCATATA	2972
Dd		188	GlnProAlaProSerThrLeuSerGlnGlnAlaLaIaTyProGlnProGlnSerGlnGln	207
OY		2973	CATGAACAACACAAACGTATACTGTAAGCAGCGCCTTACCAACTCTTTTGTCGACACAGTCT	3032
Dd		208	SerGlnGlnInThraLaTySerGlnGlnInArgPheProProProGlnGlnLeuSerGln---	226
OY		3033	GATCCTTAGGGGCACGAACTGAATTGATGTCGTGGAAAGCGCATCTCTCAACACCACT	3092
Dd		227	AspSerPheGlySerGlnAlaSer-----SerAlaProSerMetThrSerSer	242
OY		3093	AGGTCCTGGACAAAACCTCAAGAAATCATGTATAGACAGTATCAGTAACATCCAGGAGACAG	3152
Dd		243	LysGlyGlyGlnGlnAspMechAn-----LeuSerLeuGlnSerArgProSer	258
OY		3153	GCAAATMAAGATATCTGGAAGACGCTGGCTGATTTGGACAGCTCACAAAGAAGC---	3209
Dd		259	SerLeuProAspLeuSerGlySerIleAspAspLeuPrometGlyThrGlnGlyAlaLeu	278
OY		3210	GATCCTGGCCCTGCACACTGAAAAACAATTTGATTAATCTGCAGGAAGTTAAGTTCCACAG	3258
Dd		279	SerProGlyValSerThr--SerGlyIleSerSerSerGlnGly-----	292
OY		3270	GAAAACTTAAACACCAACACAG-----CGGCAACCTCGGGAAGAACAGTTGGCANA	3320
Dd		293	GluGlnSerAmProAlaGlnSerProPheSerProInThrSerProHISLeuProGly	312
OY		3321	ATTAATAAAATGAATCAATTTCTTTTCTGAAATAAGAAATTCAGTAGACSTAAATGTA	3380
Dd		313	IleArg-----GlyProSerProSerProValGlySer-----	323
OY		3381	AGCTCAGATTAACAAATAATTCAGGAGATTTAATGATGGGAGNTGCGGAGGCGAGGC	3440
Dd		324	-----ProAlaSerValAlaGlnSerArgSerGlyPro-----	334
OY		3441	GGATCTATTAATTAATCCGACGATGCGACMACTGCATAGCAGGTAAAGCCAAATCGGAG	3500
Dd		335	-----LeuSerProAlaAla-----ValProGlyAlaAn-----Gln	344
OY		3501	CTCTTATGGGCGACAAAGTTCAAGACTTTGGAAAGATGTAATGATCCA-----	3548
Dd		345	MetProProArgProProSerGlyGlnSerAspSerIleMetHisProSerMetAsnGln	364
OY		3549	GGGAGTGTATATCAATGATAGGAGGCCCTTAATAGATGTAATTAATAT-----	3596
Dd		365	SerSerIleAlaGlnAspArgGlyTyMet-----GlnAlaAsnProGlnMetPro	381
OY		3597	CAAAAACCAAGTGCACAAITGGGATCTGGAATAGGTGTTCACCTGGAACACTGCAGCT	3656
Dd		382	GlnTySerSerProGlnProGlySerAlaLeu-----	392
OY		3657	GGAGTAATATGCATATGATCATTTGTCMAAGCTCCGCGCCCCGAGATGGCAATATAGTGGSA	3716
Dd		393	-----SerProArgGlnProSerGlyGlyGlnIleHis	403
OY		3717	AGCTCTACGATATGCTAGCCTCGTTTGGCAACAACAAGTSCCAACGTCATGSGAAGCGCC	3776
Dd		404	Thr-----GlyMetGlySerTyGlnGlnAlaMetSerMetGly-----	415
OY		3777	CCAGATATGTCTAAGGAAGTTTAAATCAADAATAGCCGAACCCATTCACATCAAGGGGA	3836
Dd		416	-----SerTyGlyProGlnGlyGly	422
OY		3837	GTTGCTCAATGAGAGTGTCGAAGATTCMACATCAATTTTTCGAAGAAGCCTTAAGGGG	3896
Dd		423	-----GlnTyGlyProGlnGly	428
OY		3897	GGCAAGCCGACACAAGTACTGGAACCTGTATATCACACAAGCAAACCCCTTT-----	3950
Dd		429	GlyTyProAlaArgGlnProAsnTyArgAlaLeuProAlaAlaAsnTyProSerAlaGly	448

QY	3951	-----GGATCTGCTGGAACCTGTTAAACACCCAGGTCCGA	3986
Db	449	MetAlaGlyGlyIleAsnProMetGlyAlaGlyGlyGln	461
QY	3987	CCCGTCGACAGGTCCACCT-----CGTCCTTACACATCCATCCAGAGATCTGCCTCAGTA	4040
Db	462	---MethIAsGlyGlnProGlyIleProProIleGlyThrIleuPro---ProGlyArgMet	479
QY	4041	CCAAATAGCCACTCAATCGCCCAATCCCTCGAGTCGAAACATCTATCTCTCCG-----	4094
Db	480	SerHisAlaSerMetGlyAsnArgProTyrGlyProAsnMetAlaAsnMetProProGln	499
QY	4095	-----TACCCGCGACAAACCGCA	4112
Db	500	ValGlySerGlyMetCysProProProGlyIleMetAsnArgIleSthrgInGlnThrAla	519
QY	4113	GCAGTCANAGGGATTTGGCCGACCAACTCT-----CCAGACATG	4148
Db	520	ValAlaMetHisValAlaAlaAsnSerIleGlnAsnArgProProGlyTyrProAsnMet	539
QY	4149	GAT-----GGAACAGGGA-----TCATTATCT	4169
Db	540	AsnGlnGlyGlyMetMetGlyThrGlyProProTyrGlyGlnGlyIleAsnSerMetAla	559
QY	4170	GGATCTGTTCCGCAAGCTAATATCTTCACGCGTTCAGGACGACACACACAGTCTCTCA	4229
Db	560	GlyMetIleAsnProGlnGlyProProTyrSerMetGlyGlyThr-----Met	575
QY	4230	GCAAAACAGAACTGTTTTCAGGACACACCCA-----	4262
Db	576	AlaAsnAsnSerAlaGlyMetAlaAlaSerProGluMetMetGlyLeuGlyAspValIle	595
QY	4263	---TCGCGCGTCAAAATCAAAATCGAGTAAATCCGATCCGATCCGATCGTTCATCCGAT	4319
Db	596	LeuThrProAlaThrIlePylMetAsnAsnIleAlaAspGlyThrProIleSthrgIleuSerIle	615
QY	4320	AACCTTAAGCAGCAACCAAGTACCCCTTATCTCATCTATCCCAAGAAATTTGAGCT	4379
Db	616	SerIleSlySerSerSerSerThrThrThrAsnGlnIleThrIleSlyGlnIleu	635
QY	4380	TTCCGTCAGTCTCTGCTGCTGATG-----TTATATTTGTTTAATTTTTTAAAGAC	4430
Db	636	GlyIleGlyIleuProGluArgIleMetTyrValAspArgTyrIleu---AlaPheThrGlnIleu	654
QY	4431	AAA-----TCAATATGCAATTTGGCTTAATAATAG-----	4460
Db	655	LysAlaMetGlyMetThrAsnIleuProAlaValGlyArgIleProLeuAspIleuTyrArg	674
QY	4461	TTATATATATCA-----TAACTCGGAATTTGTATGAAAAAATCAGAAATAGAAAAAT	4514
Db	675	LeuIleTyrAlaSerValIleSlyGlnIleGlyIleuThrGlnIleValAsnIleAsnIleTyr	694
QY	4515	AAATATATTTTCCGACCGCCCATCATCTTCTTGAAATCCAAATTTCTGAGATGTTTAGA	4574
Db	695	Arg-----	695
QY	4575	GATATATCTACTATTAATAATTAAACAGAAAAATCATATCCGTTAATTGAAATCACTATT	4634
Db	696	---GluIleuAlaThrAsnLeuAsnValIleGlyThrSer-SerSerAlaIaIaSerIleu--	713
QY	4635	GTTTAATATAGCAAAATTAATAATATGCTTTATTAAT-----	4669
Db	714	---LysIleSlnIleIleGlnCysIleu-IleAlaPheGlnIleCysIleIleGlnArgIleGly	732
QY	4670	-----ATTCTCAAGCTGATAACATGAAGAAT-----A	4697
Db	732	IuAspProProProAspIlePheIleAlaIaIaAspSerIleSlySerGlnProIleSlyIleG	752
QY	4698	GGGACACCAAGCCCA-----CAGGTCGACGGGTCCACGATAAATATGTC	4739
Db	752	ImProProSerProAlaGlySerIleSerMetGlnGlyProGlnIleProGlnIleSerIle	772
QY	4740	TAAATAGAGCAAAATGAAGATGACGATTTGCTGTCATCAGTCTCGATTTTAAACCCGAT-	4798

[illegible]

```

Db      1061 rAlAlAtHrGlnArGrPrroAlaGlyGyrProGlnAaSnGlnPheProheGlnPheG1 1081
Qy      5610 TAAcCTCAATGGGTGCTGTAGGCAATGGAGTGGCCAAATATGCCACCACCATCAGCAGCGA 5669
Db      1081 YArGAsPArGyAlSeRAlA-----ProProGlyThyrAaAl 1093
Qy      5670 CGGrATCTCTGGATATGCCAGATTTAGTGGCGGAGCA-----5706
Db      1093 aGlnGlnAaMMePrProGlnMeTcGlyGyrProILeGlnAlaSeRAlaGluValAl 1113
Qy      5707 -----GGAGCCGAGAGGTATCTATGAATTCCTCCGAGAGCA 5744
Db      1113 aGlnGlnGlyThrMeTTrpGlnGlyArGAsAaPMeTThrTyAaSnTyAlaAaPArGAl 1133
Qy      5745 ACACCAaAACAAGATCACAAACAATCTGTGGGCAAGCAATGTATTAATCTTTTCAGAA 5804
Db      1133 nSerThrGlySeRAlaPrroGlnGlyPro--AlaTyhNIGlyValAaSn-----Ar 1149
Qy      5805 TTGCAATCAATATGCTATTTGTTCAGAA-----GAGGGTGAGATTACCCGCGCA 5852
Db      1149 gThrAsPGLuMeTLeuNlSthrAsPGLnArGAlaAaSnIGlnGlySerTrpProSeRhl 1169
Qy      5853 TGACGGATCAATGAATATTGGTCAACCATGTATGATTAAGGGCATGCGTCACATGCCAT 5912
Db      1169 g-----1169
Qy      5913 GCGGCAATGTAAATGGGTGCGCGGATGCCACC 5946
Db      1170 -----GlyThrArGlnPrroPro 1175

RESULT 11
US-09-270-767-61668
/ Sequence 61668, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270.767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 61668
/ LENGTH: 43
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-270-767-61668

Alignment Scores:
Pred. No.: 1,11e-13 Length: 43
Score: 236.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.96% Indels: 0
DB: 4 Gaps: 0

US-10-664-859-1 (1-6909) x US-09-270-767-61668 (1-43)
Qy      2808 AAAAACCTTAAAGATTAAACAAGTTACAGGGCAGCAATTCGCGGTATGCCATGGATA 2867
Db      1 LysaSnPrroLulYsILeAaSnYLeGlnArGrhIsaSeRValIGlyMeTProTrpILe 20
Qy      2868 GGCATGGGCGAGTGTGACTAATCTCTTATCTGTAGCCAAATTAACAACAAGCAG 2927
Db      21 GLyMeTGLyGlnValIGlyLeuThrProPrroAaSnPrroValAlaYsILeThrGlnGln 40
Qy      2928 CCACATACA 2936
Db      41 PrroISThr 43

RESULT 12
US-09-949-016-7421
/ Sequence 7421, Application US/09949016

```


Db 495 ArgProThrSerArgLeuSnaRgLeuProGluLeuGluLeuAlaIleAspAengIn 514
 QY 4278 AATCGTAGTAAATACCGGA-----TCGCAAGCGTCTTACGAT 4319
 Db 515 PheGlyGlnProGlyThrGlyAspGlnIleProTPrThAsnAsnThrValThrAlaIle 534
 QY 4320 AACTTAGCAGCAACCCA-----AGTACCCTTATCTCATTA-----TCC 4361
 Db 535 AangInSerIySerGlnAspGlnCysIleSerSerGlnLeuAspGlnLeuLeuCySpro 554
 QY 4362 CCAAGGAATTTGAGCTTTCGTCAGTCTCTGCTGATGATGATTAATTTGTTAATTT 4421
 Db 555 ProThrThrValGlnGluValArgAsnAspGluLysAlaLeuLeuGlnLeuValSerPhe 574
 QY 4422 TTT-----AAAGCAAAATCAATATGAAATTCGTTAATTAATTAATTAATTAATTA 4475
 Db 575 LeuSerGlyLysAspGlnThrGlnLeuAlaGluLeuAspArgAlaLeuGlyIleAspLys 594
 QY 4476 CTC-----GGAATTTGATGAAAAAATCAGGAATAGA----- 4508
 Db 595 LeuValGlnGlyValGlyLeuAspValLeuSerGluArgPheProProGlnGlnAlaThr 614
 QY 4509 -----AAAAATTAATTTTTC-----GGA 4529
 Db 615 ProProLeuIleMetGlnGluArgProAsnLeuTySerGlnProTySerSerProSer 634
 QY 4530 CCGCCATCCATTTCTTGAATCCAAATTTCTGAGATGATTGATGATTAATCTACTATTA 4589
 Db 635 ProThrAlaAsnLeuProSerProPheGlnGlyMetValArgGlnLysProSerLeuGly 654
 QY 4590 AAATTAACACGAAATTCATATCGTTAATGAAAAATCACTATGTTTAATGAATTT 4649
 Db 655 ThrMetProValGlnVal--ThrPro----- 662
 QY 4650 AAAAATATGTTTATTAATTTCTACAGTGTATTAACATGAAGATAG----- 4699
 Db 663 -----ProArgGlyAlaPheSerProGlyMetGlnProArgGlnThrLeuA 680
 QY 4700 --CGACCA-----AGCCCA-----CAGGTC 4718
 Db 680 snbRgProProAlaIaIaProAsnGlnLeuArgLeuGlnLeuGlnArgLeuGlnGlyG 700
 QY 4719 AGCGTCACCAAGTAAATAGTCTAATAGAGCAATAAAGATGTA-----CGATTTG 4769
 Db 700 InGln-----GlnLeuIleHieGlnAsnArgGlnAlaIleLeuAengInPheA 716
 QY 4770 CTGCAATCCAGTCT-----GCTTTAACCCGATCCACATATGCAAGC-----AATTCA 4820
 Db 716 IaAlaThrAlaProValGlyIleAsnMetArgSerGlyMetGlnGlnIleThrProG 736
 QY 4821 ATTCAACATTTAAAGCCTAATAAATGGCTCTACCAATATACAGATGAGGTAAATATTT 4880
 Db 736 InProProLeuAsnAlaGlnMetLeuAlaGlnArgGlnLeu----- 751
 QY 4881 AAATATTTTATTTAAGCTTTTGTGTTAATTAATCTTTTTCAGCGTCACATCAGC 4940
 Db 752 -----TySerGlnGlnIleArgG 758
 QY 4941 GCAAGTGCATCGGTACAAATTTAGTCGGCGCTCCGATTAATTTCCGTTAAATCCCAATAG 5000
 Db 758 In-ArgGlnLeu----- 761
 QY 5001 TGGCAATCGGCCCCACCAACAAGATGACCAAAATTCGATCAATCTCTTTGGC 5060
 Db 762 -----IleGlnGlnGlnArgAlaMetLeuMetArgGlnGlnInsPheGly 776
 QY 5061 ACAATATGCCCAACAATAACAAGTTGCTGTCACAGATGGTATGTCACGCCGAATCGG 5120
 Db 777 AsnAsnLeuProProSerSerGlyLeuProValGln-MetGly----- 790
 QY 5121 TGGTATGACATATGCGGGGTCGCCGACCGTCGACATCAATATTTGACATGGAATTAAT 5180
 Db 790 ----- 790

QY 5181 TTCGGACTAGATGATCAGAAATAGATACATAATCAAAATAATCTGATTCATTAAGA 5240
 Db 790 ----- 790
 QY 5241 TGTCTAATGAACATAAGGATCCCGAATGCTGAATCTTAATATGTGCTGACAGCGG 5300
 Db 791 -----AsnProArgLeu----- 794
 QY 5301 TCCAAATGACCGCCT-----GGCTTAATCTTAATTC 5333
 Db 795 -ProGlnGlyAlaProGlnGlnPheProTyProProAsnTyrglyThrAsnProGlyTh 814
 QY 5334 CCCCAGTGTGATTAAGAGAAATTCATAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5393
 Db 814 rProProAlaSerThrSerProPheSerGlnLeuAlaIaAsnProGluAlaSerLeuAl 834
 QY 5394 AAATTTCAAGGGGTTTTCACCTGTCGCCAGAAATGATGGGTCCAAATGCCAGTCAATT 5453
 Db 834 AsnAsnArgAsnSerMetValSerArgly-----MetThrGlyAsnIleGlyGlyGlnP 852
 QY 5454 TGTTCGAATTTCAATCCGATATTCAGGTAAAGCGAGTACCCCAACACCATACATA 5513
 Db 852 eGlyThrGlyIleAsnProGlnMetGln-----GlnAsnValPheGlnTy 867
 QY 5514 C-----ATGCCAGTAAAGGCACAGAACGCCCAACAACATTAACAATAG 5558
 Db 867 rProGlyAlaGlyMetValProGlnGlyValAlaAsnPheAlaProSerLeuSerPro 887
 QY 5559 ACCTAATATATGCGCAATG-----CCACTAGTCTGGAATTTTTCAGAGC----- 5604
 Db 887 ySerSerMetValProMetProIleProProProGlnSerLeuLeuGlnGlnIleThrP 907
 QY 5605 -----TAGCGTACCT-----CAATGGGCTGTAGGCA 5636
 Db 907 oProAlaSerGlyTyrglnSerProAspMetLysAlaTrpGlnGlnGlyAlaIleLys 927
 QY 5637 TGGTCGCCAATATAGCCACATCAGCAGCAGCAGCTCTCTGGAATGCCAGATTC-- 5694
 Db 927 nAsnAsnValPheSerGlnAlaValGlnAsnGlnProThrProAlaGlnProGlyValTy 947
 QY 5695 -----ATGCCGGACACAGAGCCCGAGATGCTAATGA 5729
 Db 947 rAsnAsnMetSerIleThrValSerMetAlaGlyIleAsnThrAsnValGlnAsnMetAs 967
 QY 5730 TTCTTCGGAGAGCAACACCAAGATCAAAACAATCTCTGGGGCAAGC----- 5781
 Db 967 nProMetMetAlaGlnMetGlnMetSerSerLeuGlnMetProGlyMetAsnThrValCy 987
 QY 5782 -----AATGTTAATTAATCTTTTCAGAATTG 5807
 Db 987 sProGlnGlnIleAsnAspProAlaLeuArgIleThrGlyLeu-----Tyrcy 1003
 QY 5808 CATCAAAATGCTATTTGTTGAC-----GAAGAGGTGATTAACCGCGCATACAGG 5858
 Db 1003 sAsnGlnLeuSerSerThrAspLeuLeuTyThrGluAla-----AspGln 1018
 QY 5859 ATCAATGAATATTTGTCACCAATCTATG-----ATAAG 5891
 Db 1018 yThrGlnGlnValGlnGlnValGlnValPheAlaAspValGlnCysThrValAsnLeuVa 1038
 QY 5892 GGGCATGCTGCATGCGATCCAGCGCCAAATGTAATGGTGGCGGATCCACCGTTAA 5951
 Db 1038 lGlyGlyAspProTyLeuAsnGlnProGlyProLeuGlyThrGlnLysProThrSerGln 1058
 QY 5952 CAGGCAATTCAGTTTGACAG 5973
 Db 1058 yProGlnThrProGlnAlaGln 1065

RESULT 13
 US-09-125-635-4
 ; Sequence 4, Application US/09125635
 ; Patent No. 6562589


```

; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1420
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-125-635-4

Alignment Scores:
Pred. No.: 1.81e-11 Length: 1420
Score: 222.00 Matches: 267
Percent Similarity: 30.04% Conservative: 150
Best Local Similarity: 19.24% Mismatches: 496
Query Match: 1.84% Indels: 475
DB: Gaps: 65

US-10-664-859-1 (1-6909) x US-09-125-635-4 (1-1420)

QY 3120 GTAGACAGTTCATCATCCAGGAGTCCACAGCAATTAAGATCTGACAGCCTGCG 3179
DB 95 ValGlnLysAlaAspValSerSerThrGlyGlnGlyValIleAspLysAspSerLeuGly 114
QY 3180 GTTGGATTGGGACAGCTCACAAAGAGATCTGCTGACAACTGCAAAACAACTT 3239
DB 115 ProLeuLeuLeuGlnIleAlaLeuAspGlyPheLeuPheValAlaAsnArgAspGlyAsnIle 134
QY 3240 GTATCACTGCAGAGAGTTAAAGTTCCAGAGAAACCTTACACCAACAAGCGCAACAT 3299
DB 135 ValPheVal-----SerGlnAsnValThr--GlnTyrLeuGlnTyr 147
QY 3300 CGGGAAGAACGTTGGCAAAATTAATAATTAATTTCTTTCTGAAATAGAG 3359
DB 148 LysGlnGlnLysPheValAlaAsnThrSerValTyrAsn--IleLeuHISGlnGlnAspArg 166
QY 3360 AATTCACTAGACGTAATGTAAAGCTCAGACATAACAATAATCCAGAGAGATTAAATGATG 3419
DB 167 LysAspPhe-----LeuLysAsnLeuProLysSerThrValAsn 179
QY 3420 GGGATGTCGGTGCGGAGGCGGATCTATTATTAATCCAGATGCGACAACCTGCATATG 3479
DB 180 GlyValSerTyr-----ThrAsnGlnThrGlnArgGlnLysSerHis 193
QY 3480 CCAAGTAAACCCAAATCGAGCTCTTATCGCGCAACAAAGTTCAGAGACTTTCGGAAGATGA 3539
DB 194 ThrPheAsnCysArg-----MetLeuMetLysThr----- 203
QY 3540 ATGCATCCAGAGGATGTTATATGATATGATGCGCTTAATAGATTAATTAATCA 3599
DB 204 -----ProHisAspIleLeuGlnAspIleAsnAlaSerProGlnMetArgGlnAspTyr 221
QY 3600 AAAACACAGTGTGCAATGTGATCTGAGTAGGTGTCTCACTGCACAACTGCAGCTGGA 3659
DB 222 GlnThr----- 223
QY 3660 GTAATATCATATATGATCTCAAGCTCCGGGCCCCGGAATGGAATATATGAGGAGC 3719
DB 224 -----MetGlnCysPheAlaLeuSerGlnProArgAla--MetMetGlnGln 238
QY 3720 TCTACGGATATGCTAGCCTGTTGGACACAACAAGCTCAACGTC----- 3764
DB 239 GlyGlnAspLeu-----GlnSerCysMetIleCysValAlaAspArg 252
QY 3765 -----ATCGGAACGGCCCCCAT 3782
DB 253 IleThrThrGlnGlnArgThrPheProSerAsnProGlnSerPheIleThrArgHisAsp 272

```

```

QY 3783 ATGCTTAAGAGATTAAATAGATAGCCGAAAC-----CATTCACATCAAGGGGCA 3836
DB 273 LeuSerGlnLysValValAlaValAsnIleAsnIleAsnThrAsnSerLeuArgSerMetArgProGly 292
QY 3837 GTTGCTCAATG-----GAG 3851
DB 293 PheGlnAspIleIleArgArgCysIleGlnArgPhePheSerLeuAsnAspGlnSer 312
QY 3852 TGGTCGAAGATTAAATCAATTTTTCGAAGAACGCTCAAGGGGGCAAGCCACAGCA 3911
DB 313 TrpSerGlnLysArgHis--TyrGlnGlnValTyrLeuAsnGlnLysAlaGlnThrPro 331
QY 3912 GTC-----ACTGGAAGCTGTAGTCAACAAGCAAGCAAGCCCTCTCGA 3953
DB 332 ValTyrArgPheSerLeuAlaAspGlnThrIleVal--ThrAlaGlnThrLysSerLys 350
QY 3954 TCTGTGGAATACTGTTAAACAACAG----- 3980
DB 351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeuGln 370
QY 3981 -----GTGCAACCCCTG-----CAAGTCCACCTCTCTCT 4010
DB 371 ArgGlnGlnAsnGlyTyrArgProAsnProAsnProValGlnGlnGlyIleArgProPro 390
QY 4011 TACCACCTCCATCCAGAGATCTGCGTCAATGACCAATGACCAATCGCCCAAT----- 4064
DB 391 MetAlaGlyCys-----AsnSerSerValGlyGlyMetSerMetSerProAsnGlnGly 408
QY 4065 -----CCCTGAGTCCCAACAATCTATCTCCCGTACCGCGGACCAACCGCAGCA 4115
DB 409 LeuGlnMetProSerSer--ArgAlaTyrGlnLysLeuAlaAspProSerThrThrGln 427
QY 4116 GTCATGGAGTTGCC-----ACCACTCTCTGAC 4145
DB 428 MetSerGlyAlaArgTyrGlnLysSerSerAsnIleAlaSerLeuThrProGlyProGly 447
QY 4146 ATGATGGAACAAGAGATCTATCTGATCTGTCGCCAAGCTAAATCTTGACGGTTACG 4205
DB 448 MetGlnSerProSerSerTyrGlnAsnAsnValTyrGlyLeuAsnMetSerSerProPro 467
QY 4206 GCAGGACAACAACAGTCTCTCAGCAACAAGAACTGTTTTCAGGACAGACCCCATCG 4265
DB 468 HisGlnSerProLysLeu-----Ala 474
QY 4266 CCGTCAATCAAAAT-----CGTAGCAAAATACCGGATCGTCAACGCTTTACG 4316
DB 475 ProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerProLysIleAlaSer 494
QY 4317 CATACCTTAAGC-----AGCAACCCAAAGTACCCCTTATCTCATCTATCCCAAGAA 4370
DB 495 HisGlnPheSerProValAlaGlyValHisSerProMetAlaSerSerGlyAsnThrGly 514
QY 4371 TTGAGTCTTTTCGTCAGTCTCTGCTGTATGTTATTTGTTAT----- 4418
DB 515 AsnHisSerPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGlnGlyValGly 534
QY 4419 -----TTTTTAAAGCAAAATCAATATG 4442
DB 535 ThrSerLeuSerThrLeuSerSerProGlyProLysLeuAspAsnSerProAsnMet 554
QY 4443 AATTGGCTTAATTAATAGTTATATATTAACATACTCGAAATTTGATAGAAATAATCAG 4502
DB 555 AsnIleThrGlnPro-----SerLysValSer 563
QY 4503 AATGAAAAAATAATTAATTTTCCGACCGCCCATCATTTCTTG-----AATCCA 4553
DB 564 AsnGlnAspSerLys-----SerProLeuArgPheTyrCysAspGlnAsnPro 579
QY 4554 ATTTCTGAGTAGTGTGTT-----AGAGATAAATCTACTATTTAAA----- 4592
DB 580 ValGlnSerSerMetCysGlnSerAsnSerArgAspHisLeuSerAspLysGlnSerLys 599

```

```

QY 4593 -----TTAAACGAAATTCATATC 4613
Db 600 GUserSerValGluGlyAlaGluAenGlnArgGlyProLeuGlnSerGlyHisLeu 619
QY 4614 CGTTAATTGAAATCACTATTGTTTAAAGAAATTAATAATATGTTTATTAATATTT 4673
Db 620 Lys-----Leu 621
QY 4674 CTACAGGTATACATGAAAGTAGGCGACCAAGCCAGAGGTGACGGTCAACGATA 4733
Db 622 LeuGlnLeuLeuThrCysSerSerAspAparGlyHis-SerSerLeuThrAsnSer-- 640
QY 4734 ATAGCTTAATAGAGCAAAATTAAGATGACGATTGCTGCATCCAGTCCTGTTTAACC 4793
Db 641 -----Pr 641
QY 4794 CGCATCCAGATATGCAAGCAATTCAAATTGACATTAAACGCTTAATAATGGCTGA 4853
Db 641 CLeuAspSerSerCysLeuGlnSerSerValSerValThrSerPro--SerGlyValSe 660
QY 4854 CCAATATACAGATGAGGT-----AAATATTTAAATATTTATTTATTA 4895
Db 660 rSerSerThrSerGlyGlyValSerSerThrSerAsnMetHisGlySerLeuLeuGlnG 680
QY 4896 CGTTTGTGTTAATTATCTTTTTCAGCGT----- 4929
Db 680 uLysHisArgGlyLeuHisLeuLeuGlnAenGlyAsnSerProAlaGluValAlaLys 700
QY 4930 -----CAGCATCAGCGCAAGGTGATCGGATTCGATTAAGTCGGCGTCCGATTA 4979
Db 700 sIleThrAlaGluAlaIleThrGlyLysAspThrSerSerIleThrSerCysGlyAspGlyAs 720
QY 4980 TATTCGCTAAATCCCAATAGTGGCAATCGGCGCCACCAACCAAGATGACCCAAACTT 5039
Db 720 nValValLysGlnGlnGlnLeuSerProLysLysGluAsnAsnAlaLeuLeuArgLys 740
QY 5040 C-----GATCCATCTCTTCTTTGCCA----- 5061
Db 740 rLeuLeuAspArgAspAspProSerAspAlaLeuSerLysGluLeuGlnProGlnValAlaG 760
QY 5062 -----CAATGTCCCAACACACTAACAGTTGGCTGCCAGATGGGTAGTGC 5108
Db 760 uGlyValAspAsnLysMetSerGlnCysThrSerSerThrIleProSerSerSerGlnG 780
QY 5109 AGCCGAACTGCTGTATGACGATGTGGGGGTCCGGACCGTCCGACATCAATATTA 5168
Db 780 uLysAspProLysIleLysThrGlnThrSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 800
QY 5169 GCATGGAATTAATTTCCGAGCTAGATGATCAGGAATA-----GATACCATTAATCAAA 5222
Db 800 uAspAlaIleLeuGlyAspLeuThrSerSerAspPheTyrAsnAsnSerIleSerSerAs 820
QY 5223 TAACTGCAATTCATGAATGTCGTAATGAACCTAATGGGTCCCGAATGCTGAATCTTA 5282
Db 820 nGlySerHis-----LeuGlyThrLys 827
QY 5283 AATGTCGACGAGCGGCTCCAAATGA----- 5310
Db 827 eGlnGlnValPheGlnGlnGlyThrAsnSerLeuGlyLeuLysSerSerGlnSerValGlnSe 847
QY 5311 -----CCGCCCT-----GGCTTTAATCTTAATTCCTCCCAATGAGTGATTAAGAGA 5354
Db 847 rIleAspProProLysAsnArgAlaValSerLeuAspSerPro----- 861
QY 5355 GAATTCATAGGTGTGGCTGTGGCTCAGCAACTCTTCAAACTTTCAAGGGGTGTGTC 5414
Db 862 -ValSerValGlySerSerProProValLysAsnIleSerAlaPhePro-----MetLe 879
QY 5415 ACCTGTGCCAGATAGGTGGAATGCCA-----GTCAATTTTG 5456
Db 879 uProLysGlnProMetLeuGlnGlyLysAsnProArgMetMetAspSerGlnGlnLysnArg 899
QY 5457 TTCCGAATTC-----AATCCGAATATTCAGGTAAGGCGAGTACCCCAACACCAT 5507

```

```

Db 899 ySerSerMetGlyGlyProAsnArgAsnValThrVal---ThrGlnThrProSerSerG 918
QY 5508 ACAATAC---ATGCCA-----GTAAAGGCGACGAGCAACCAACAA 5546
Db 918 yAspThrGlyLeuProAsnSerLysAlaGlyArgMetGluProMetAsnSerAsnSerMet 938
QY 5547 TAACAACAATGAGAGCTAAT---AATGTGCAATGACCACTAGCTGGAATTTTTCGAGAG 5603
Db 938 rGlyArgProGlyArgLysPheTyrAsnThrSerLeuPro----- 950
QY 5604 GTACGCTAACCTCTCAATGGGTGCTGTAGGCAATGGGTGCCAATATGCCACATCAGC 5663
Db 951 -----ArgProAlaLeuGly-----GlySerIleProThrLeuProLeuArg 964
QY 5664 CAGCGAGGTACTCCTCGAATGCCAGATTGATGGCGGAGCAGAGAGCCGAGGTATGCT 5723
Db 964 gSerAsnSerIleProGlyAlaArgProValLeu----- 975
QY 5724 AATGAATTTCTCCGAGAGCAACACCAGAAAGATCACAAACATCCTGG----- 5775
Db 976 -----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 990
QY 5776 -CAAGCAATGGTATTAATCTTTCAGAAATTCGAATCAAAATGCTATTGTGACCAAGA 5834
Db 990 oMetGlyMetGlyAlaAsnProTyr-----GlyGlnAlaAlaAlaSerAsnGlnLe 1007
QY 5835 GGGTGATTAACCGGCGCATGAGCATGAAT----- 5868
Db 1007 uGlySerThrPro-----AspGlyMetLeuSerMetGlnGlnValSerHisGlyThrG 1025
QY 5869 -----ATTGTCAACATCTATGATTAAG 5891
Db 1025 nAsnArgProLeuLeuArgAsnSerLeuAspLeuValGlyProProSerAsnLeuG 1045
QY 5892 GGGCATG-----CGTCCACATGCCACCGGCCAAATGTAT 5927
Db 1045 uGlyGlnSerAspGluArgAlaLeuLeuAspGlnLeuHisThrLeuLeuSerAsnThrAs 1065
QY 5928 GGGTGGCGAGGACCCCGTAAACGCAATT----- 5961
Db 1065 palAthrGlyLeuGlnGlnIleAspArgAlaLeuGlyIleProGluLeuValAsnGlnG 1085
QY 5962 -----CAGTTTGACAGTCATCGATGATGATTAAGTGTGTCGGAGA 6002
Db 1085 yGlnAlaLeuGlnProLysGlnAspAlaPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1105
QY 6003 TCCGTCATATTTTTCATCAAGCTTCCGCAACAGCGTGCACACATGTTTGATC 6062
Db 1105 pGlnLysAlaGlyLeuTyrGlyGlnThrTyrProAlaGlnGlyProProMetGlnGlyG 1125
QY 6063 AGCACAA---CAGGCCAATCAGCTTAAGACACACATCAATGAACCTAGTGAAT 6119
Db 1125 yPheHisLeuGlnGlnGlnGlnSerProSerPheAsnSerMetMetAsn----- 1140
QY 6120 GTGTCAAAACAATCGGACTTGTCAGTGGCACAGGG----- 6156
Db 1141 -GlnMetAsnGlnGlnGlnGlnAsnPheProLeuGlnGlnMetHisProArgAlaAsnIleMe 1160
QY 6157 -----CAGATCCAGTCGATGGGCAAGACATGCGCAGGG 6191
Db 1160 rArgProArgThrAsnThrProLysGlnLeuArgMetGlnLeuGlnGlnArgLeuGlnG 1180
QY 6192 TCAGTCTTTAATTTGACCTTAATTAATTAATTA-----ATGTCACTGC 6236
Db 1180 yGlnGlnPheLeuAsnGlnSerArgGlnAlaLeuGlnLeuLysMetGluAsnProThrAl 1200
QY 6237 CGAAAGTGCAGTGCT-----ACTAAAGGTGTCTTGCAATCA 6275
Db 1200 acGlyAlaAlaAlaValMetArgProMetMetGlnProGlnGlnGlnPheLeuAsnAlaG 1220
QY 6276 TTTTCGAGGTCCCTCTTCAAGACTG----- 6303

```



```

Db      873  LAAGINGInIleAsnAlaGlnArgProGluProGlnSerAspAsnAspSerSerAlaThr 893
QY      3428 GGGTGGCGAGGCGGAGTCTATTATAATCCGACGATCGCAACATCGATATGACAGTAA 3487
Db      893  ICysSerAlaAspGluAspValAspGluProGluArgGlnArgMetPheProMetAs 913
QY      3488 CGGCAATCGGAGCTCTTATCGGCGACAAAGTTGAGACTTTCCGAA----- 3533
Db      913  pSerLysProSerLeuAsnProThrGlySerIleLeuValSerSerProLeuLysPr 933
QY      3534 -----GATGTAATGCATCCAGGGGAGTGTATATACATATGAGGTGC 3574
Db      933  oAsnProLeuAspLeuProGlnLeuGlnHisArgAlaAlaValIleProPheMetValSe 953
QY      3575 CGTAATAGAGTGTATATATATCAAAAAACAGTGTGCATGTGCATGTGCAGTACGATGTG 3634
Db      953  ICysThrProCysAsn----- 11 959
QY      3635 TGTCACTGGAACAACGTGCAGCTGAGTA----- 3662
Db      959  eProIleGlyThrProValSerGlyTyrAlaLeuTyrGlnArgHisIleLysAlaMetHis 979
QY      3663 -----AATGTAATATGATTCCTC 3682
Db      979  sGlnSerAlaLeuLeuGlnGluGlnArgGlnArgGlnGlnGlnIleAspLeuGluCysArg 999
QY      3683 AAGCTCGGCGCCCGCAATGCAATATGATGGAAGCTCAGCATATGCTAGCTCGT 3742
Db      999  gSerSerThrSerPro----- 1004
QY      3743 TGGCAACACAAGCTGCACGTCATCGGAACGCCCCAGATATGTCTAAGGAAGTTTAA 3802
Db      1005 -----CysGlyThrSerLysSerProAsnArg----- 1013
QY      3803 TCAAGATAGCGGAACCATTCACATCAAGGGGAGTTGCTCAATGAGTGTGCAAGAT 3862
Db      1014 -----GluTyrPheValIle----- 1018
QY      3863 TCATCAATTTTTCGAAGAA-----CGCTCAAGGGGG 3898
Db      1018 uGlnProAlaProHisGlnValIleThrAsnLeuProGluGlyValArgLeuProThrHis 1038
QY      3899 CAAGCCGACAGACATCTGGAACGTGATGACCAACAAGCAAC----- 3944
Db      1038 rArgProThrArgProProProProLeuIleProSerSerLysThrValAlaSerG 1058
QY      3945 -----CCTTCTGAGTCTGTGTAACCTGTTAAACAACAGGTGCGACCCCTGCAAGTGC 4000
Db      1058 uLysProSerPheIleMetGlyLysSerIleSer-----GlnGlyHis 1072
QY      4001 ACCTCCTCCTTACCATCTCATCCAGAGATCTGCTACATCAATAGCCATCATGCGC 4060
Db      1072 rProGlyThrTyrLeuThrSerHisAsnGlnAlaSer-----TyrThrGlnGluThrPr 1090
QY      4061 CAATCCTCG-----AGTCCAAACA 4081
Db      1090 oLysProSerValGlySerIleSerLeuGlyLeuProArgGlnGlnIleValLysSe 1110
QY      4082 TCATCTCTCCCG-----TCACGGCGG----- 4103
Db      1110 rAlaThrLeuProTyrIleLysGlnGluPheSerProArgSerGlnAsnSerGlnPr 1130
QY      4104 -----ACAAACCCGACGAGT 4117
Db      1130 oGlnGlyLeuLeuValArgAlaGlnHisGlnGlyValValArgGlyThrAlaGlyAlaI 1150
QY      4118 CATG-----GGATTGCCGACCACTCTCCAGCATGATGGAACAGG 4159
Db      1150 eGlnGlnGlySerIleThrArgGlyThrProThrSerLysIleSerValGluSerIlePr 1170
QY      4160 ATCATTATCTGATCTGTCCGCAAGCTATATCTTCACGTTTCAGCAGGACAAACAAC 4219

```

```

Db      1170 oSerLeuArgLysSerIleThrGlnGlyThrProAlaLeuProGlnThrGlyIleProTh 1190
QY      4220 AGTGCCTCAGCAAAACAAGAACTGTTTTCAGGACAGACCCCATCGCCGTCAAAATCAAA 4279
Db      1190 r----- 1190
QY      4280 TCGTAGTGAATATCCCGATCGTCAAGCTTTTACGCAATTAACATTAAAGCAGAACCAAG 4339
Db      1191 -GluAlaLeuValLysGlySerIleSerArgMet----- 1201
QY      4340 TACCCCTTATCTCATCTATCCCAAGGAATTGAGTCTTCGCGACGTCTGCTGG 4399
Db      1202 ---ProIleGluAspSerSerProGluLysGlyArgGlnGluAlaIleSerLysGlyHis 1220
QY      4400 TATGTTATATTTGTTTAAATTTTAAAGACAATCAATATAGAAATGGCTTAATATAA 4459
Db      1220 sValIleTyr-----GluClyLysSer----- 1227
QY      4460 GTTATATATATACATACTCGAAATTTGATA-----GAAAAATTCAGAAATAGAAAAA 4513
Db      1228 -----GlyHisIleLeuSerTyrPheAsnIleLysAsnAlaArgG 1241
QY      4514 TAAATTTATTTTC-----CGGACGCCCATCATTTCTTGAAATCCAAATTCGAGTGAATTG 4569
Db      1241 uGlyThrArgSerProArgThrAlaHis----- 1250
QY      4570 TTAGAGATATCTACTATTAAATTAACAACGAAATTCATATCCGTTAATTGAATAATCA 4629
Db      1251 -----GluIleSerLeuLysArgSerTyrGluSerValGluLys----- 1263
QY      4630 CTATGTTTAAATGAATAATTAATAATATGTTTATATATATTTCTACAGGTGATACAT 4689
Db      1264 -----AsnIleLysGlnGlyMetSerMet 1271
QY      4690 GAAAGTAGCGACCAAGCCACAGGGTCAAGCGTCAACGATTAATGCTAATAGAGGC 4749
Db      1271 rArg-----GluSerProValSerAlaProLeuGluG 1282
QY      4750 AAATGAAGATGATGCAATTTGTCATCGACTCTGTGTTTAAACCGCATCCACATATGCA 4809
Db      1282 y-----LeuIleCysArgAlaLeuProArgGlySerProHisSerAspLeuLys 1298
QY      4810 AAGCAAT-----TCAAATTCAGATTAACCGCTTAATAAGGCTGTACCA----- 4857
Db      1298 sGluArgThrValLeuSerGlySerIleMetGlnGlyThrProArgAlaIleThrHisLys 1318
QY      4858 -TATACAGATGAGGATTAATTTAAATTTTATTTAAGTTTGTGTTAATTTATCT 4916
Db      1318 rPheGluAspGlyLeuLysTyrProLysGlnIle----- 1329
QY      4917 TCTTTTTCAGCGTCAGCA-----TCAGCGCAAGGTGATCCGTAACAATTTAGTGC 4967
Db      1330 -----LysAspGluSerProProIleArgAlaPheGlnGlyAlaIleThrLysGly 1347
QY      4968 GCGCTCCGATATATTTCCGTTAAATCCCAATAGTGCATCG-----CCGCG 5015
Db      1347 sProTyrAspGlyIleThrThrIleLysGluMetGlyArgSerIleHisGluIleProArg 1367
QY      5016 ACCAAACAAGATGACCAAAACCTTCATCCAAATCTCTTTCGCAACAATGCCACA 5075
Db      1367 gGlnAspIleLeuThrGlnGluSerArgLysThrProGluValAlaGlnSerThrArgPr 1387
QY      5076 ACTAACAGTTCGTCAGATGAGTATGCA-----GCCGGA-- 5115
Db      1387 oIleIleGlnGlySerIleSerGlnGlyThrProIleLysPheAspAsnAsnSerGlyG 1407
QY      5116 -----ACTGTTGATACATGATGAGGGG 5141
Db      1407 nSerAlaIleLysHisAsnValLysSerLeuIleThrGlyProSerLysLeuSerArgG 1427
QY      5142 TCCGGACCGTCCGACATATATGACGATGGAATTAATTCGAGACTAGATGATCAGG 5201
Db      1427 yMetProProLeuGluIle-----ValProGluAsnIleLysValValGluArgGlyLysTyr 1446

```


QY 4729 ----- 4729
 Db 661 rgInThrLeuAsnArgProProAlaIaIaProAsnGlnLeuArgLeuGlnLeuGlnIa 681
 QY 4730 -----GTAATAGTCTAATAGAGCAATAAAGATGTA-----CGATTGG 4769
 Db 681 rgleuGlnGlnGlnGlnGlnLeuIleHleGlnAsnArgGlnAlaIleLeuAsnGlnIa 701
 QY 4770 CTGCATCCAGTCTCT---GGTTTAAACCCGATCCACATATGCAAGC-----AATTCAA 4820
 Db 701 IaIaIaThrAlaProValGlyIleAsnMetArgSerGlyMetGlnGlnIleThrProG 721
 QY 4821 ATTCAGCATTTAAACGCTTAATAAATGGCTCTACCAATATACAGATGAGGTAATATT 4880
 Db 721 InProProLeuAsnAlaGlnMetLeuAlaGlnArgGlnLeu----- 736
 QY 4881 AAATATTATTATTAACTTTTGTGTATTATTATCTTTTTCAGGCTCAGCATCAGC 4940
 Db 737 -----TyrSerGlnGlnHisArg 743
 QY 4941 GCAAGTGSATCCGTACATTTAGTGGGGCTCCGATTAATATCCGTAATCCCATAG 5000
 Db 743 In-ArgGlnLeu-IleGlnGlnAlaMet-----LeuMetArgGlnGlnSerPh 760
 QY 5001 TGGCAATCGGCGCCCAACCAAGATGACCAGAACTTCGATTCATCTTCTTGGC 5060
 Db 760 eGlyAsnAsnLeuProPro----- 766
 QY 5061 ACAATGTCCCAACCAATCAAGTGTGTCAGATGGGTAGTCAGCCGAGACTGG 5120
 Db 767 -----SerSerGlyLeuProValGlnThrGln 775
 QY 5121 TGGTATGACGATGATGGGGGGTCCGGGACCGTCCGATCATGATTTAGCATGAATAT 5180
 Db 775 Y----- 775
 QY 5181 TTCGGAGCTAGATGATCAGAAATAGATACATTAATCAAAATTAATCTGATTCATGA 5240
 Db 775 ----- 775
 QY 5241 TGTGTATGTAACACTCAATGGGTCGCCGATGCTGAATCTTAATAATGCGTAGAGGG 5300
 Db 776 -----AsnProArgLeu----- 779
 QY 5301 TCCAAATGACCGCT-----GGCTTAATCTTAATTC 5333
 Db 780 -ProGlnGlyAlaProGlnInPheProTyProProAsnTyGlyThrAsnProGlyTh 799
 QY 5334 CCCCATGTGTGATTAGAGAAATTCATAGGGTGTGGCTGTGGCTCAGCAAACTCTTC 5393
 Db 799 rProProAlaSerThrSerProPheSerGlnLeuAlaIaAsnProGlnAlaSerLeuAl 819
 QY 5394 AAACCTTCAAGGGTGTTCACCTGTGTCAGAAATGATGCGCAATGCCAGTCAATTT 5453
 Db 819 aAsnArgAsnSerMetValSerArgGly-----MetThrGlyAsnIleGlyGlnPh 837
 QY 5454 TGGTTCGAATTTCAATCCGATATTCAGGTAAGGCGAGTACCCCAACACCATACATA 5513
 Db 837 eGlyThrGlyIleAsnProGlnMetGln-----GlnAsnValPheGlnTy 852
 QY 5514 C-----ATGCCAGTAAGGACAGAACGCCCAACCAATTAACAACATG 5558
 Db 852 rProGlyAlaGlyMetValProGlnGlyGlnAlaAsnPheAlaProSerLeuSerProGln 872
 QY 5559 AGCTAATATGTGGAATG-----CAACTAGTCTGGAATTTTGCAGAGG----- 5604
 Db 872 ySerSerMetValProMetProIleProProGlnSerSerLeuGlnGlnInThrPr 892
 QY 5605 -----TAGCTAACCT-----CAATGGGTGCTGTAGAGCAA 5636
 Db 892 oProAlaSerGlyTyGlnSerProAspMetLysAlaTrpGlnGlnGlnAlaIleGlyAs 912
 QY 5637 TGGGTCCCAATATGCCACCATCAGCAGGAGGTACTCTGGAATGCCAGGATTG-- 5694

Db 912 naAsnAsnValPheSerGlnAlaValGlnAsnGlnProThrProAlaGlnProGlyValTy 932
 QY 5695 -----ATGGCGGACCAAGAGCCGAGCTTGCTAATGA 5729
 Db 932 rAsnAsnMetSerIleThrValSerMetAlaGlyGlyAsnThrAsnValGlnAsnMetAs 952
 QY 5730 TTCTTCGGAGAGCAACACAGAAACAAGATCACAAACAAATCTGGGGCAAGC----- 5781
 Db 952 nProMetMetAlaGlnMetGlnMetSerSerLeuGlnMetProGlyMetAsnThrValCy 972
 QY 5782 -----AATGTATTAACTTTCTTCAGAAATG 5807
 Db 972 sProGlnGlnIleAsnAspProAlaLeuArgHisThrGlyLeu-----TyrCy 988
 QY 5808 CAATCAAAATGCTATTGTTGAC-----GAAGAGGGTGGATTACCCGCCATGACGG 5858
 Db 988 sAsnGlnLeuSerSerThrAspLeuLeuTyThrGlnAla-----AspGln 1003
 QY 5859 ATCAATGAATATTGTGCAACCATCTATG-----ATPAG 5891
 Db 1003 yThrGlnGlnValGlnGlnValGlnValPheAlaAspValGlnCysThrValAsnLeuVa 1023
 QY 5892 GGGCATGCGTCCACATGCCATGGCGCCAAATGTAATGGGTGCGCGATGCCACCGTTAA 5951
 Db 1023 lGlyGlyAspProTyLeuAsnGlnProGlyProLeuGlyThrGlnTyProThrSerGln 1043
 QY 5952 CAGCAAAATTCAGTTTGACAG 5973
 Db 1043 yProGlnThrProGlnAlaGln 1050

Search completed: November 2, 2005, 22:23:59
 Job time : 275 secs

THIS PAGE BLANK (USPTO)

Db 221 ----- 221
Qy 2208 TTAAGTCGAGCATCAGACAGATTTCATTTTGGGTGATATATAGAAGTGCCTC 2267
Db 221 ----- 221
Qy 2268 ACACTCTTCTGGCGCTTCACACTAGTAGAGTCCGCCGAGATATTATATAGAT 2327
Db 221 ----- 221
Qy 2328 GATTACAGATTATTTATTTTATGATTTATTAATAAATCTTATTATTCAATT 2387
Db 221 ----- 221
Qy 2388 TACACTTAATAATTGAAAGATTTCAAAGACAGTACACGGAAGAAAAAGATCGTCC 2447
Db 222 -Pro -Vallyslilegialarglileserasnapselthrthrglulyslyser 240
Qy 2448 TTGCAATGAATAATGACGAATAGAGATGGAAGGTGCAATCAGTTGAATCCGATTTT 2507
Db 241 Leuthrmetasnapselgumetsermetcylgylcyasnglnleuasnproasphe 260
Qy 2508 ATCAATGAATCTTTAAATATCTGCAATTTCAGCATATTAGTAAGCGAGTAGACA 2567
Db 261 lileasnliuserleuasnproalilaserlelleuvalserglyvalglypro 280
Qy 2568 ATACCCGGGAATCGAGATTGAGCGGAGACGGAAATTTATGACTGCACACCGCAATGA 2627
Db 281 lleprolylileglyvalglyalaglythrclyasnleuethrlalaasnliaasnily 300
Qy 2628 ATCTCCCTGGGTAGCAGTAATTGTTGATTAATGCAACAGCAAAATTCACATTCGTG 2687
Db 301 lileserserglyseraserasnCysleuasptrymetcnglnnglnasnhsillepheval 320
Qy 2688 TTTTCACTGACGTGGCAACAAAGGGCGCAATCAGTTTAAAGCGGTCAATTTCAACT 2747
Db 321 Pheserthrghlnleualaasnlyeglyalaglyservalleuaserglylnpheglnthr 340
Qy 2748 ATTATTCGCTATCTGACACTGACCTGCTCTTAACAAAAGCTTCCTGGAAGACTTTTATG 2807
Db 341 llellelathyrthyscysrhnglnproalathrlysserphelengluasphepmet 360
Qy 2808 AAAAACCTTTAAAGATTAAACAAGTTACAGCGCACAAATCCGTGGATGCCATGATA 2867
Db 361 lyeasnproleuyslileasnlyseuglntrghshanserValglymetprotrpille 380
Qy 2868 GGCATGGGCGAGTTGACTTAACCTCTTAATCTGTAGCCAAATTAACACACAGAG 2927
Db 381 GlymetcylglnvalglyleuthrProproasnproValalalybillethnglnngln 400
Qy 2928 CCACATCAAAAGACCGTAGGCTTATTTGAACCCCAATTCAATCAACATGAAGACAGAA 2987
Db 401 ProthiethrlystrValglyleuethlyserProglinpheasnnglnhsigliuasnserly 420
Qy 2988 CGTAGTCTGAAGCGGCGCTAGCAACTTTTGTGACACAGTGTGATTCCTATGGGCAAC 3047
Db 421 ArgserthrValseralaproserasnserPheValaspglinsarbpromeglyasn 440
Qy 3048 GAAACTGAATGTAGTCTGGGAAGGCGAGTCTCAAAACACCACTAGGTCTGACAAAC 3107
Db 441 GluthrcluleumetCyslrpgluglyglyseraseranthrseratrgserglyGlnasn 460
Qy 3108 TCACGAATCATGTAGACAGTATCATGTATCCAGCGATGACAGGCAATTAATCACTG 3167
Db 461 SertrgnsnhsValaspserlileserthserserglinsaralilelyslileu 480
Qy 3168 GAAGCAGCTGGCTGTGATTGGAGCAGGTCAAAAGAGAGCATCTGGCGCTACAACT 3227
Db 481 GlulalalaglyValaspheuglyglnvalthrlysglyseraspproglyleuthrthr 500
Qy 3228 GAAAAACAATTTATCACTGACGAGGATTAGGTTCCAGCGAAACCTTTACACCAAA 3287

Db 501 GluasnasnilleValserleuglnglyVallyseValproaspgluasnleuthrProglin 520
Qy 3288 CAGCGGCAACATGGGAGAGACAGTTGGCAAAATATAAATGAATCAATTCCTTTT 3347
Db 521 GlharfglnhsartrggluglnglnleuValalylllelyblysmecasnnglnphele 540
Qy 3348 CTTGAAAATGAGATTCACTAGAGAGCTATGTAAAGTCAACAGATTAACAAAATTCAGCA 3407
Db 541 ProgluasnnglnsnserValglylaleasnValsersercnglnlethrlybilleProgly 560
Qy 3408 GATTTAATGATGGGAGATGTGGGTGGGAGCGGAGATCTATTTAATCCGACATGCGA 3467
Db 561 AspIeumetmetcylmetsercylglyglyglyserllelleasnprothmetcarg 580
Qy 3468 CAATGCATATGACAGTAACGCAATCGAGCTCTTATCGGAGCAAGTTCAGACTT 3527
Db 581 GlnleuhsimectProglyasnValalysercylleuuseralatrserSerglyleu 600
Qy 3528 TCGGAAGATGTAATGATCCAGGGAGATTATATACAGATATGGGTGCCGTAAATAGATG 3587
Db 601 SerclunspValmetchisproglyaspvalilleseraspmetcylalaValilleglyCys 620
Qy 3588 AATTAATATCAAAAACCAAGTGTGCAATGTGATCTGAGTAGTGCTTTCACTGCAACA 3647
Db 621 AsnasnasnnglnystrhservalglnCysglysercylvalglyValalithrglythr 640
Qy 3648 ACTGACGCTGAGTAATGTCATATGATGATGCTGCTCAAGCTCCGCGCCCGGAATGGCAAT 3707
Db 641 ThrAlaValaglyValasnValasnmetchisCysseaserSererglyalaproasnGlyasn 660
Qy 3708 ATGATGGGAAGCTCTACGATATGCTAGCTCGTTTGGCAACACAACTGCAACGTATC 3767
Db 661 MetmetcylserserthrtraspmetleuValaserPhehlysnthrserCysasnValille 680
Qy 3768 GGAACGCGCCCAATATGCTTAAGAAATTAAATTAAGATACGCCAATCCATTCAT 3827
Db 681 GlyThrAlaproaspmetserlysglyValleuasnnglnaspseralagtrhlsesehs 700
Qy 3828 CAAGGGGAGTTGCTCAATGAGAGTGGTGGAGATTCAATCAATTTTCAAGAAAGCG 3887
Db 701 GlnglyglyValalaglnmetcylutrpserlyslileglnhsnglnpheglnlyarg 720
Qy 3888 CTCAGAGGGGCGCAAGCCGACAGACAACTGCACTGGAATCTGATACACAAAGCAACCCCT 3947
Db 721 LeuylsglylysparoaglnvalThrglyThrValValProglnglnnglnthrPro 740
Qy 3948 TCTGATCTGTGAAAATCTGTTAAACACCAAGTGCAGCCCTGCAAGGTCCACTCT 4007
Db 741 SerglyserglylysnserleuAasnnglnValargProleuglnglyProProPro 760
Qy 4008 CTTTACCATCTCCATCCAGAGATCTGCGTCAAGTACCAATACCACTCAATGCCCAATCC 4067
Db 761 ProlyrhlsierlileglnargseralaserValProilleathrglnserProasnpro 780
Qy 4068 TCAGTCAACAACATCTATCTCCCGTCAACCGCGCAACCGCAGCAGTATGGAGATTG 4127
Db 781 SerSerProasnhsnleuuserleuProserProargthrthrlalaValillemetcylleu 800
Qy 4128 CCGAACCACTCTCTAGCATGATGAACAGATCATTAATCTGTGATCTGTTCCGAAGCT 4187
Db 801 ProthrasnserProsermetaspglyThrglyserleuuserglyserValProglinala 820
Qy 4188 AATATCTGACGGTTCAGGAGGACAAACAAGTGCTCAGCAACCAAGAACTGTTT 4247
Db 821 AsnthrserthrValglnalaglythrThrValleuserAlaasnlyspasnCysPhe 840
Qy 4248 CAGCAGACACCCCATCGCCGTCAATCAAAATCGTAGAATAACCGATCGTCAACG 4307
Db 841 GlnAlaaspthrProserProserasnnglnasnargserarganthrlyserSerSer 860
Qy 4308 GTTCTTAAGCAATACTTAAGCAGACCAACCAAGTACCCTTATCTCATCTTACCCCAAG 4367
Db 861 ValleuthrhrhsnleuuserSerasnProserThrProleuserhlsleuSerProlys 880

QY 4368 GAATTGAGTCTTCGCGTAGCTCTGCTGTGATGTTAATTGTTAATTTTTTAA 4427
 |||||
 Db 881 GluPheGluSerPheGlyGlnSerSer----- 889
 QY 4428 GACAAATCAAAATGGAATTCGCTTAATTAAGTTATATATTACATACTCGAAATTTG 4487
 ----- 889
 Db 4488 ATGAAAAAATCAGAAATAGAAAAAATTAATTATTTTCGACCGCCATCATTCTTG 4547
 889 ----- 889
 QY 4548 AATCCAAATTTCTGAGTGAATGTTAGAGATACTATCTAATTAATAATTAACGAAAAAT 4607
 889 ----- 889
 Db 4608 CATATCCGTTAATTGAAATCACTATGTTTAAATAAGAAATTAATAATATGTTATTATA 4667
 889 ----- 889
 QY 4668 ATATTTCTACAGGTGATTAACATGAAAGTAGCGACCAAGCCACAGGCTCAGCGCTCAC 4727
 |||||
 Db 890 -----Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907
 QY 4728 CAGTAATAGTCTTAATAGAGCAAAATPAAAGATGTACGATTGCTGCATCCAGTCTGATT 4787
 907 roValAsnSerLeuIleGlnIAsnLysAspValArgPheAlaIAsnSerSerProGlyP 927
 QY 4788 TTAAACCCGATCCACATATGCAAGCAATTCGAATTCAGATTTAAACGCTATTAATAATG 4847
 |||||
 Db 927 heAsnProhIAspProhIAsnMetGlnSerAsnSerAsnSerAlaLeuAsnAlaTyrLysMetG 947
 QY 4848 GCTCTACCAATATACAGATGAGAGTAATTTAATATTATTATTAAAGTTTGTGTT 4907
 947 LysSerThrAsnIleGlnMetGln--Arg----- 955
 QY 4908 AATTATCTTCTTTTCAGCGTCAAGCATCAGCCGCAAGGTGATCCGTAATTTAGTGC 4967
 |||||
 Db 956 -----GlnIAsnSerIAsnGlnGlyGlySerValGlnPheSerArg 968
 QY 4968 GCGCTCGATTAATATTCGCTAATTCGCAATAGTGCATTCGCGCCGCCCAACAAAGAT 5027
 |||||
 Db 968 gAspSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgProProProAsnLysMe 988
 QY 5028 GACCCAAAATTGATCCAACTCTTCTTTGCGCAAAATGTCCCAACAACTAACAAGTTG 5087
 |||||
 Db 988 tThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGlnIleuThrSerCy 1008
 QY 5088 CGTGTCCAGCATGGTAGTCCAGCCGGAACCTGGTGATGACGATAGGGGGTCCGG 5147
 |||||
 Db 1008 sValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMetMetGlyGlyProG 1028
 QY 5148 ACCGTCCGACATTAATTTAGCATGGAATTAATTTGGGACTGATGATCAGGAATAGA 5207
 |||||
 Db 1028 yProSerAspIleAsnIleGlnIleGlyIleIleSerGlyLeuAspGlySerIlyIleAs 1048
 QY 5208 TACCATTAATCAAAATTAACGTGCTCATTCATGAATGTCGTAATGAATCCTCATGGGTCC 5267
 |||||
 Db 1048 pThrIleAsnGlnAsnAsnCyshiSerMetAsnValValMetAsnSerMetClyProh 1068
 QY 5268 AATGCTGAATCCTAAATGTGCTAGCAGCGGCTCCAAATGACCGCTGGCTTTAATCC 5327
 |||||
 Db 1068 gMetLeuAsnProLysMetCysValAlaGlyIleProAsnGlyProProGlyPheAsnPr 1088
 QY 5328 TAATTTCCCAATGCGTGAATTAAGAGAAATTCATAGAGTCTGGCTGGCTCACAACA 5387
 |||||
 Db 1088 oAsnSerProAsnGlyGlyLeuArgGlnAsnSerIleGlySerGlyCysGlySerIlyIleAs 1108
 QY 5388 CTCTTCAAACTTTCAGAGGGGTGTCTCCACTGGTGCAGAAATGATGGTGAATGCAGT 5447
 |||||
 Db 1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetClyArgMetProVa 1128

QY 5448 CAATTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAGTACCCCAACCACT 5507
 |||||
 Db 1128 lAsnPheGlySerAsnAsnPheAsnProAsnIleGlnValLysAlaSerThrProAsnThrI 1148
 QY 5508 ACAATACATGCGAGTAAGGCGCACAGAACGCCCAACAATTAACAATAGAGCTAATTA 5567
 |||||
 Db 1148 eGlnTyrMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnGlyAlaAsnAs 1168
 QY 5568 TGTGGAATGCCACCTAGTCTGGAATTTTGGAGAGTACAGCTTAACCTCAATGGGTGC 5627
 |||||
 Db 1168 nValArgMetProProSerSerLeuGluPheLeuGlnArgTyrAlaAsnProGlnMetClyAl 1188
 QY 5628 TGTAGGCAATGGGTGCGCAATTTGCCACCATCAGCCAGCGACGCTACTCTGGAATGCC 5687
 |||||
 Db 1188 aValGlyAsnGlySerProIleCysPProProSerAlaSerAspGlyThrProGlyMetPr 1208
 QY 5688 AGGATTTGATGGGGGCGACGAGCGGAGGTATGCTAATGAATTTCTTCGAGAGCAACA 5747
 |||||
 Db 1208 oGlyLeuMetClyAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSerSerGlyGlnGlnI 1228
 QY 5748 CCAGAAACAGATCACAACAATCTGCGGCAAGCAATGCTATTAACTTCTTTCAGAAATG 5807
 |||||
 Db 1228 eGlnAsnLysIleThrAsnAsnProGlyAlaAsnSerAsnGlyIleAsnPhePheGlnAsnCy 1248
 QY 5808 CATCAATATGCTATTGTTGACGAAAGGCTGCAATTCACCGGCATGACGATCAATGA 5867
 1248 sAsnGlnMetSerIleValIAspGlyGlnGlyGlyLeuProGlyLysAspGlySerMetAs 1268
 QY 5868 TATTTGGTCAACCATGATATGTAAGGGGCAATGGGTCCACATGTCATGCGGCCAAATGTAT 5927
 |||||
 Db 1268 nIleGlyGlnProSerMetIleArgGlyMetArgProhIAsnMetArgProAsnValMe 1288
 QY 5928 GGGTGGCGGATGCCACCCGTTAAGAGCAATTCAGTTTGGACAGTCTCGATGATGAT 5987
 |||||
 Db 1288 tGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGlnSerSerAspGlyI 1308
 QY 5988 TGACTGTGCGGGATCCGTCATCATTTTTCATAACGCTTCTCGAACAAGCGCTGGAGC 6047
 |||||
 Db 1308 eAspCysValGlyAspProSerSerPhePheThrAsnIAsnSerCysAsnSerAlaGlyPr 1328
 QY 6048 ACACATGTTTGGATCAGCAACAAGGCCAATCAGCTTAAGACAAACATTAAGAAGAT 6107
 |||||
 Db 1328 ohiAsnMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGlnIAsnIleLysAsnI 1348
 QY 6108 ACTTAGGAATGTGTCAAAACCAATCGGACTTGCAGTGGCAAGGCGAGATCCAACT 6167
 |||||
 Db 1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnGlyGlnIleGlnMe 1368
 QY 6168 GCATGGGCAAGACATGCCAGGGTCACTTTAATTGACCTTACTATTAATTAATTAT 6227
 |||||
 Db 1368 uhiAspGlyGlnGlyLysAlaGlnGlyGlnSerLeuIleGlyProThrAsnAsnAsnLeuMe 1388
 QY 6228 GTCAATCGCGGAAGTGCAGTGTACTTAACGGTGTCTGGCATCAATTTGCTAGTGC 6287
 |||||
 Db 1388 tSerThrAlaGlySerValSerAlaThrAsnGlyAlaSerGlyIleAsnPheValGlyPr 1408
 QY 6288 CTCTTTCAGGACCTGAAGTATGCCAGCAATATCAATGTTTTCAGCAGCAGTATATGC 6347
 |||||
 Db 1408 oSerSerThrAspLeuLysTyrAlaGlnGlnTyrhisSerPheGlnGlnGlnLeuTyrAl 1428
 QY 6348 TACCAACACCAAGATCAACAACAACAGCATATGCCACGACGACACAGCAACATGAT 6407
 |||||
 Db 1428 aThrAsnThrArgSerGlnGlnGlnGlnIAsnMetCysGlnGlnIAsnIleGlnSerAsnMetI 1448
 QY 6408 AACCAATGCGCGGAATTTATCAACCAATGCAAGTCTTGTCAACAA 6456
 |||||
 Db 1448 eThrMetProProAsnLeuSerProAsnProThrPhePheValAsnLys 1464
 RESULT 2
 ABW01527
 ID ABW01527 standard; protein; 1464 AA.
 XX
 AC ABW01527;

XX 15-JAN-2004 (first entry)
 DT Drosophila species legless (lgs) protein.
 DE
 XX Legless protein; lgs; cell fate disorder; blood disease; gene therapy;
 KW cancer; tissue regeneration; tissue repair; cytostatic.
 XX
 OS Drosophila sp.
 XX US2003114413-A1.
 XX
 PD 19-JUN-2003.
 XX
 PE 19-DEC-2002; 2002US-00322579.
 XX
 PR 28-JUL-2000; 2000US-0221502P.
 PR 27-JUL-2001; 2001US-00915543.
 XX
 PA (UYZU-) UNIV ZURICH.
 XX
 PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
 XX WPI; 2003-829432/77.
 DR N-PSDB; AAD62641.
 XX
 PT Novel lgs polypeptide useful for isolation of lgs-binding proteins,
 PT diagnosing disorders of cell fate, treating diseases such as cancer.
 XX
 PS Claim 5; Fig 2; 0pp; English.
 XX
 CC The invention relates to novel legless (lgs) proteins and polynucleotides
 CC encoding such proteins. lgs sequences are useful for the treatment of
 CC disorders of cell fate such as differentiation or proliferation. The
 CC invention is used to treat blood disease or a cancerous condition
 CC characterized by over-stimulation of the Wnt pathway such as colon,
 CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and
 CC is administered to prevent progression from a pre-neoplastic or non-
 CC malignant condition to a neoplastic or malignant state. It is
 CC administered to promote tissue regeneration and repair. The invention is
 CC also useful in the therapy of diseases cost by an over-activation of Wg
 CC pathway. It is useful for reducing lgs gene expression in an invertebrate
 CC or vertebrate organism or an invertebrate or vertebrate cell line. The
 CC invention is also useful in gene therapy. The present sequence is
 CC Drosophila species legless (lgs) protein
 CC
 XX
 SQ Sequence 1464 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 1464
 Score: 7257.00 Matches: 1460
 Percent Similarity: 73.01% Conservative: 1
 Best Local Similarity: 72.96% Mismatches: 3
 Query Match: 60.29% Indels: 537
 DB: 7 Gaps: 5
 US-10-664-859-1 (1-6909) x ABM01527 (1-1464)
 QY 468 ATGCCCGGAGTCCCAACCAACAGCCGCAACCAACTCCGATGCTCTCTCAACAGT 527
 DB 1 MetProtrSerProtrGlnGlnGlnProGlnProbsenSerAaPaLaSerSerThSer 20
 QY 528 GCATTCGATCAATCTCTGAGCAGCAGTGGAAATGGGAGCTCGGGGAGAGAGT 587
 DB 21 AlaSerIysSerAsnProGlyAlaAlaIleGlyAsnGlyAspSerAlaAlaSerArgSer 40
 QY 588 TCTCCGAAGACCTTAATAGGAACTCTTTTACTTGTGCGCGGGTAAGACTTGATT 647
 DB 41 SerProIysThrIleuAsnSerGluProPheSerThrIleu-Ser----- 54
 QY 648 GAATTCCTCTTGTTCGGAATTATTAACAATTCTGTGCTTTCAGATCAATAAATTGAC 707
 DB 55 -----ProAspGlnIleIysLeuTh 61

QY 708 GCCAAGAAAGGAGCTGAGAAAAGCGGACTATCACTAGTAAAGCTGCACTGGAG 767
 DB 61 rProGluGluGlyThrGluYsSerGlyLeuSerThrSerAspYsAlaAlaThrGlyG 81
 QY 768 AGCCCGAGGAGTGAATAATATCTGCCGAGGAGCAAACTATGCTAAGGAGACTTAC 827
 DB 81 yAlaProGlySerGlyAsnMetLeuProGluGlyGlnThrMetLeuArgGlnAsnSerTh 101
 QY 828 GAGCAATATCACTCGTCTAGTGGCTTCTCCCAAAAATCCAGTGAACACTCGAATAG 887
 DB 101 rSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSerGlnIleSerAsnSe 121
 QY 888 CAGCAATGTGCTGCTACAGTGGGCTTACTCAGATGTGATGTTGAGCAGCAATGAA 947
 DB 121 rSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCysAspGluGlnSerI 141
 QY 948 GAAAAACAAATGTAGTGTGAAGAGCAGAGAGCTGTGAAGACTGCGCTACAAATGCTTTA 1007
 DB 141 slyAsnIysCysSerValIysAspGluGlnAla----- 152
 QY 1008 AAATTTAAATGATATTGGCGCTTACCTTTGTTAATCATTTAATTGTTTTTTCCTA 1067
 DB 152 ----- 152
 QY 1068 TACTTAATTTTATGTTTAACTTGTAACCTTGACATAAAGCTCGGAGCTCGATCAA 1127
 DB 152 ----- 152
 QY 1128 AACGACATTTCTGTGAAACCGTAATTAAGCTCATATAAATTAATTCATCTTGATGA 1187
 DB 152 ----- 152
 QY 1188 ATGCATATCATATGATGTACTCAAAACATCTGAAGAAAGCCTCAATTTGATCACTAATT 1247
 DB 152 ----- 152
 QY 1248 AGTTTGAAAAAATTGCTGTACTTTAAGATATATTAATTAATAAATTGGCTGAGTGA 1307
 DB 152 ----- 152
 QY 1308 AATGATATATATGATCAATAATTTTATTAACCTGAAGCAATTTGAAATAGCCGTG 1367
 DB 152 ----- 152
 QY 1368 CTACGAGATGCTACTAGACCGGCTGAAGGCTAATTTTATTAAGCTGTCTAAT 1427
 DB 152 ----- 152
 QY 1428 ATTCCATTAACATTAAATGTCCTATTCAGAAATTAAGTTCTAATTAAGCAAAAGTCAAGC 1487
 DB 153 -----GluIleSerSerAsnIysAlaIysGlyGlnAl 163
 QY 1488 AGCTGTGGCGGCTGCGAAGACAGTTCTACATCCAGTTGACTGTCAAGAGAACCCAC 1547
 DB 163 aaIaGlyGlyGlyCysGluThrGlySerThrSerSerIleThrValIysGluIupProth 183
 QY 1548 CGATGCTTATAGGAGATTAGTAATATGAAAAAGAAAGAAAGAAATATATTCGCCAAC 1607
 DB 183 rAspValIleuGlySerIleuValAsnMetIysGlyGluIuArgIuAsnIleSerProth 203
 QY 1608 GATGCCCCCTGTTGTTTGGTCAATTGGTATGACAGAGCAACCTCGCTCACCGGG 1667
 DB 203 rMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAsnSerAlaThr----- 221
 QY 1668 TAAGTTTAAAGATCATATAAAGCAAAATTAACAAGAAATTAATGTCAGTTACCAATTTTA 1727
 DB 221 ----- 221
 QY 1728 TTGATATGCAAGAACTACTATAGGATATCTCTGCTTTAATTTAATTTAATTAG 1787
 DB 221 ----- 221

QY	1788	GAAATACGAATATTCTCAATTTGTAAAAATAAAATGATTAATTACATAGAAATTTAAAAAC	1847
Db	221	-----	221
QY	1848	CTTTGAAATTAGACATACCTCTCCAAAATCAGTAATCATTTGGAAACGAGAGTGATC	1907
Db	221	-----	221
QY	1908	CGAAGGAGACTACTATATAAACCTTTTGAGCTATCTGATCTACTGACCGTACTAAAAATGA	1967
Db	221	-----	221
QY	1968	TTAGTTAGAAAAATGGGTGTAATTTGTAGAAAGTTTTCATTTTAGAAGAAATGTGATT	2027
Db	221	-----	221
QY	2028	ATTTTATTAAACCCCTTCAGCGGAACTACATTTGTTCAGATAATTTTGGAAAAACAA	2087
Db	221	-----	221
QY	2088	TGGTTAAAGTTGGAAGATGCCCTATATAAACAGAAATTCACCGTTTCAAAATACTAACAGGTT	2147
Db	221	-----	221
QY	2148	TTTGATTTAATTTTGATTAATAATGAGAAATTATCACACTTCAGTTAAAAATGTTTAATTCGA	2207
Db	221	-----	221
QY	2208	TTAAGTGCAGCAATCACAGAGATTTCCATTTTTCGTGTATATATAGAAAGTGCCTTC	2267
Db	221	-----	221
QY	2268	ACACTCTTCTGGCGGCTTCACCACTACGTGAGTTCCGCCGAGATTTATATATAGAT	2327
Db	221	-----	221
QY	2328	GATTTACAGATTAATTAATTTTATATGTGATTTTAATAATATCTTATTATTCATT	2387
Db	221	-----	221
QY	2388	TACATAGTTAAAAATTGAAGAATTTCAACGACAGTACCAGCGAAAAAAGATCGTCC	2447
Db	222	-Pro--ValIleIleGluArgIleSerIleAspSerThrThrGluValGlyGlySerSer	240
QY	2448	TTGCAATGAAATATATGACGAAATGAGCATGGAAGGCTCAATTCAGTTGAATCCGATTTT	2507
Db	241	LeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGlnLeuAsnProAspIle	260
QY	2508	ATCAATGAAATCTTTAAATATATCTCTGCATTTTCGACATATTAATGTAAGCGAGTAGACCA	2567
Db	261	IleAsnGlnSerIleuAsnAsnProAlaIleSerSerIleLeuValSerGlyValGlyPro	280
QY	2568	ATACCCGGAATCGAGATTGGAGCGGGGACGGGAAATTTATGACTGCGCAACGCCAATGGA	2627
Db	261	IleProGlyIleGlyValGlyAlaGlyIleThrGlyAsnLeuLeuThrAlaAsnAlaAsnGly	300
QY	2628	ATCTCTCGGTAGCAGTATTTGTTGGATTATACGACAACAGCAAAATCAGATATTGCTG	2687
Db	301	IleSerSerGlySerSerIleAsnAspIleuMetGlnGlnIleAsnHisIleIlePheVal	320
QY	2688	TTTTCACTCAGCTGGCCCAACAAGGGCCGGAATCATGTTTAAAGCGGTCAATTTCAACT	2747
Db	331	PheSerThrGlnLeuAlaAsnIleGlyIleGlnSerValIleuSerGlyGlnPheGlnThr	340
QY	2748	ATTATTGGGTATCACTGCACTCAGCGCTGCTCAAAAAGCTTCCTGGAAAGACTTTTATG	2807
Db	341	IleIleValIleIleCysThrGlnProAlaThrIleuSerPheLeuGluAspPheMet	360
QY	2808	AAAAACCCCTTTAAATATTAACAAGTTTACAGCGGCACAAATTCCTCGGTATCCATGATTA	2867
Db	361	LysAsnProIleuLeuIleAsnIleuLeuGlnArgHisAsnSerValGlyMetProIleIle	380
QY	2868	GGCATGGGGCAGGTGGACTTAATCTCTCTTAATCCTGTAGCCAAAATTAACAACAACAGAG	2927

[illegible]

Db 741 SerGlySerGlyGlyLeuSerLeuAsnAsnGlnValArgProLeuGlnGlyProProPro 760
Qy 4008 CCTTACCATTCCATCCAGAGATCTGGCTGACTACCAATACCACTCAATGGCCCAATCCC 4067
Db 761 ProYHisSerLeuLeuGlnArgSerAlaSerValProLeuAlaThrGlnSerProAsnPro 780
Qy 4068 TCGAGTCGCAAAACATGATCTCTCCCGTCACCGGGGCAACCGGAGAGTCAGGGATTG 4127
Db 781 SerSerProAsnAsnLeuSerLeuProSerProArgThrAlaAlaValMetGlyLeu 800
Qy 4128 CCGACCAACTCTCTAGCATGATGAGAACAGATCATATATCTGATCTGTTCCGCAAGT 4187
Db 801 ProThrAsnSerProSerMetArgGlyThrGlySerLeuSerGlySerValProGlnAla 820
Qy 4188 AATATCTTCCAGCGTTCAAGGAGGACCAACAACAGTGTCTCAGCAACAAGACTGTTT 4247
Db 821 AsnThrSerThrValGlnAlaGlyThrThrValLeuSerAlaAsnLysAsnCysPhe 840
Qy 4248 CAGGACAGACACCCATCGCCGTCAAAATCAAAATCGTAGTAATAATACCGGATCGTCAAGC 4307
Db 841 GlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsnThrGlySerSerSer 860
Qy 4308 GTTCTTAACGAACTTAACAGCAACCAACCAAGTACCCCTTATCTCATCTATCCCAAG 4367
Db 861 ValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSerHisLeuSerProLys 880
Qy 4368 GAATTTGAGCTTTTCCGTCAGTCTCTCGCTGATGATGTAATTTGTTAATTTTTTAAA 4427
Db 881 GluPheGluSerPheGlyGlnSerSer----- 889
Qy 4428 GACAAATCAAAATGAAATTCGTTAATTAATGATTATATATATACATTAACTCGGAATTTG 4487
Db 889 ----- 889
Qy 4488 ATGAAAAAATCAGAAATAGAAAAAATTAATTTTCCGACCGCCCATCTTCTTG 4547
Db 889 ----- 889
Qy 4548 AATCAATTTCTGAGATGATTGTAGAGATTAATCTACTATTAATAATTAAACAGAAATTT 4607
Db 889 ----- 889
Qy 4608 CATATCCGTTAATGAAAAATCAGTATGTTTAATTAAGAAATTAATAATATGTTTATTATA 4667
Db 889 ----- 889
Qy 4668 ATATTTCTACAGGTGATTAACATGAAAAAGTAGGAGCAACCAACGAGGTCAAGCGGTAC 4727
Db 890 -----Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907
Qy 4728 CAGTAATAGTCTTAATAGAGGCAAAATTAAGATGATGCAATTTGCTGCATCCAGTCTGGTT 4787
Db 907 roValAsnSerLeuLeuGlnAlaAsnLysAspValArgPheAlaAlaSerSerProLysP 927
Qy 4788 TTAACCGGATCCATATGATCAAGCAAAATCAAAATTCAGATTAAACGCTATTAATAATG 4847
Db 927 heAsnProHisAspProHisMetGlnSerAsnSerAsnSerAlaLeuAsnAlaIlyLysMetG 947
Qy 4848 GCTCTACCAATATACAGATGAGAGTAATATTTAAATATTTTATTAAACGTTTTGTGTT 4907
Db 947 LysThrAsnLeuLeuGlnMetGlu--Arg----- 955
Qy 4908 AATTTATCTTTCTTTTCAAGGTCAAGATCAAGCGCAAGGTGATCCGTACATTTTATGCG 4967
Db 956 -----GlnAlaSerAlaGlnGlyLysValGlnPheSerArg 968
Qy 4968 GCGCTCGATATATTCCTGTAATCCCATATGAGCAATCGGCGCGGACCAAAACAAGAT 5027
Db 968 gaYgSerAspAsnLysProLeuAsnProAsnSerGlyAsnArgProProProAsnLysMet 988
Qy 5028 GACCCAAAACCTTGATCAATCTCTTTTGGCAAAATGTCCCAACAACATAACAGTTG 5087
Db 988 cThrGlnAsnPheAspProLysSerSerLeuAlaGlnMetSerGlnGlnLeuThrSerCys 1008

Qy 5088 CGNGTCACAGATGGGTAGTCCAGCGGAACTGGTGTATGACATGATGGGGGCTCCGG 5147
Db 1008 sValSerSerMetGlySerProLysGlyThrGlyLysMetThrMetMetGlyLysProG 1028
Qy 5148 ACCGTCGACATCAATATTTAGCATGAATTAATTTCCGGACTAGATGATCAGAAATAGA 5207
Db 1028 yProSerAspLysAsnLysGlnHisGlyLysLysLysSerGlyLeuAspGlySerLysLysLe 1048
Qy 5208 TACCATTAATCAAAATTAATCTGATCAATGAAATGTCATGATCAATCAATGGGTCCCG 5267
Db 1048 pThrLysGlnAsnGlnAsnAsnCysHisSerMetAsnValAlaMetAsnSerMetGlyProAr 1068
Qy 5268 AATGCTGAATCCCAAAATGCGGTACAGCGGCTCCCAATGACCCCGCTTAAATCC 5327
Db 1068 gMetLeuAsnProLysMetCysValAlaGlyLysProAsnGlyProProLysPheAsnPr 1088
Qy 5328 TAATCCCGCAATGATGATTAAGAGAAATTCATAGAGGTCTGGCTGTGCTCAGCAAA 5387
Db 1088 oAsnSerProAsnGlyGlyLeuAspArgLysAsnSerLysGlySerCysGlySerAlaAs 1108
Qy 5388 CTCTTCAAACTTTCAAGGGGTTGTTCCACTGTGTCAGAAATGATGGGTGAAATGCCAGT 5447
Db 1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetGlyArgMetProVa 1128
Qy 5448 CAATTTGTTGGAATTTTCAATCCGAATTAATTCAGTTAAAGGGGAGTAAACCCCAACCAT 5507
Db 1128 LAsnPheGlySerAsnPheAsnProAsnLysGlnValLysValaSerThrProAsnThrL 1148
Qy 5508 ACAATACATGCCAGTAAGGAGCAGAACGCCCAACAACAATTAACAATGAGACTATAA 5567
Db 1148 eGlnLysMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnGlyLysAlaAsnAs 1168
Qy 5568 TGTGCAATGCCACTAGTCTGGAATTTTGGCAGAGTACCTTAACCTCAATGGGTGC 5627
Db 1168 nValArgMetProProSerLeuGlnPheGlnArgLysAlaAsnProGlnMetGlyAl 1188
Qy 5628 TGTAGGCAATGGGTGCGCAATATGCGCACCATAGCCAGGAGGAGTACTCTGGAATGCG 5687
Db 1188 aValGlyAsnGlySerProLysProLysProSerAlaSerAspGlyThrProLysMetPr 1208
Qy 5688 AGGATTGATGGCGGAGCAGAGAGCCGAGGTATGTAATGAATTTCTCCGAGAGCAACA 5747
Db 1208 oGlyLeuMetAlaGlyProGlyAlaGlyLysMetLeuMetAsnSerSerCysLysGlnHis 1228
Qy 5748 CCAGAACAAGATCAACAACAATCTGCGGCAAGCAATGATTAATCTTTCAGAATG 5807
Db 1228 eGlnAsnLysLysLeuThrAsnAsnProGlyAlaSerAsnGlyLysAsnPheGlnAsnC 1248
Qy 5808 CATCAAAATGCTATGTTGAGCAAGAGGTGATTAACCGGCGCATAGAGGATCAATGAA 5867
Db 1248 sAsnGlnMetSerLysValAspGlyLysGlyLysLeuProGlyHisAspLysMetCAs 1268
Qy 5868 TATTGTCACAACATCTATGATTAAGGGGATGCTGCCATGCCATCGGCAATATGAAT 5927
Db 1268 nLysGlyGlnProSerMetLysArgLysMetArgProHisAlaMetArgProAsnValMe 1288
Qy 5928 GGGTGGCGGATGCCACCGGTTAACAGGCAAAATTCAGTTTGCACAGTCAATCGGATGTAT 5987
Db 1288 cGlyAlaArgMetProProValAsnArgLysLysGlnPheAlaGlnLysSerSerAspGlyL 1308
Qy 5988 TGACGTGTGCGGGAGCCGTCATCATTTTCACTAAGCTCTTGCAACAGCGCTGAGCC 6047
Db 1308 eAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCysAsnSerAlaGlyPr 1328
Qy 6048 ACACATGTTTGGATCAGCAACAAGGCCAATCAGCTTAAGACAAACATAAAGACAT 6107
Db 1328 oHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGlnHisLysLysAsnL 1348
Qy 6108 ACTTAGTGAAATGTGCAAAACAATCGGAGCTTGCAGATGAGCAAGGGGCAATCCACT 6167
Db 1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnGlyLysLysLysLys 1368


```
QY 6168 GCATGGGACAGACATGCGAGGTCAGTCTTTAATTGACCTACTAATAATTAATTAAAT 6227
DB 1368 UHISGLYGLINGLYHIALAGLNGLYNLSERILEUILLGLYPROTHRAASHANLEUWKE 1388
QY 6228 GTCAACTGCGGAGAGTGTCAAGTCTACTAAGCGTGTCTGTGGATCAATTTGCTAGTCC 6287
DB 1388 TSEHTHRLAAGLYSERVALSERIALATHRANGLYVALSERGLYLALASHPHEVALGILYPR 1408
QY 6288 CTCTTACGACGACCTGAAGATGCGCAGACATATCATAGTTTTCAGACAGAGTTATATGC 6347
DB 1408 OSESRERTHRAPELLELYETRLAAGLNGLYNLYRHSERPHGGLNGLNGLEUURYRXL 1428
QY 6348 TACCAACACCCAGAGTGTACACACACATATGACCGACGACCCAGACCAACATGAT 6407
DB 1428 ATTHRAENTHRAARGSERGLNGLNGLNGLNHISMECHLSGLNGLNGLNSERANMELIIL 1448
QY 6408 AACCAATGCGCGGATTTTATCAACCAATCCAGCTTCTTGTGCAAAA 6456
DB 1448 ETHMETPROFROASHLEUSERPROASHPROTHRPHEPHEVALASHNLYS 1464
RESULT 3
ADJ71911
ID ADJ71911 standard; protein; 1464 AA.
AC ADJ71911;
XX
XX 20-MAY-2004 (first entry)
XX
DE Fruit fly legless (lgs) polypeptide.
XX
XX Fruit fly; legless; lgs; cell differentiation disorder;
XX cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;
XX breast; head; neck; brain; thyroid; skin; blood disease;
XX tissue regeneration; tissue repair; cytoskeletal.
XX
XX Drosophila melanogaster.
XX
XX US2004038901-A1.
XX
XX 26-FEB-2004.
XX
XX 22-SEP-2003; 2003US-00664859.
XX
XX 28-JUL-2000; 2000US-0221502P.
XX
XX 27-JUL-2001; 2001US-00915543.
XX
XX (UYZU-) UNIV ZURICH.
XX
XX
XX Basler K, Brunner E, Froesech B, Kramps T, Peter O;
XX
XX WPI; 2004-203288/19.
XX
XX N-PSDB; ADJ71889.
XX
XX Novel polypeptide sharing one or more homologue amino acid domains with
XX legless protein being functional homologue of legless, useful for
XX diagnosing disorders of cell fate.
XX
XX Example 2; SEQ ID NO 1; 62pp; English.
XX
XX The invention relates to a polypeptide sharing one or more homologous
XX amino acid domains with a legless (lgs) protein and is therefore a
XX functional homologue of lgs. The invention also relates to a nucleotide
XX sequence encoding a protein present in invertebrate and/or vertebrate
XX organisms, the nucleotide sequence encoding a protein comprising a
XX positive function in a regulatory pathway and the use of the polypeptide
XX for the isolation of lgs-binding proteins by carrying out an assay chosen
XX from an in vitro binding assay with such a peptide or a co-
XX immunoprecipitation from vertebrate or invertebrate cell lysates or a
XX mammalian or yeast two hybrid assay. The polypeptide and polynucleotide
XX are useful for treating disorders of cell fate, which involves
XX administering therapeutic compounds chosen from invertebrate and
XX vertebrate lgs protein homologues or fragments, antibodies, antibody
XX fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,
```

```
CC small peptides or chemical and natural compounds being capable of
CC interfering with lgs function, synthesis and degradation. The disorders
CC are related to cell differentiation or cell proliferation. The compound
CC is administered to treat a cancerous condition by preventing progression
CC from a pre-neoplastic or non-malignant condition to a neoplastic or
CC malignant state. The cancerous condition is characterized by over-
CC stimulation of the Wnt pathway and is medulloblastoma or cancer of the
CC colon, breast, head and neck, brain, thyroid or skin. The therapeutic
CC compound may also be administered to a blood disease to promote tissue
CC regeneration and repair. This sequence represents the Drosophila lgs
CC polypeptide of the invention.
XX
XX Sequence 1464 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 1464
Score: 7257.00 Matches: 1460
Percent Similarity: 73.01% Conservative: 1
Best Local Similarity: 72.96% Mismatches: 3
Query Match: 60.29% Indels: 537
DB: 8 Gaps: 5
US-10-664-859-1 (1-6909) x ADJ71911 (1-1464)
QY 468 ATGCCCCGAGTTCACACCAACACGCGCAACCAACTCCGATGCTCTCCACAACT 527
DB 1 METPROARGSERPROTHRGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGLNGL 20
QY 528 GCATCTGATCAAAATCTCGAGACGAGTCCGAAATGGGAGCTCGCGCGGCGAGCAAGT 587
DB 21 ALASERGLYSERASHPROGLYALALALIEGLYASNGLYASPSERIALALASERAGSER 40
QY 588 TCTCCGAGAACCTTAAATAGCAACCCCTTTCTACTTTGTCCCGGGTAAAGCTGTATT 647
DB 41 SERPROLYSERTHLEUASNSERGLUPROPHESERTHLEU-SER----- 54
QY 648 GATTTCTCTTGTCCGAAATTAACAACTTTCTGTGTTTCCAGATCAATAATTAATTGC 707
DB 55 -----PROASPGINLILELYSEUTH 61
QY 708 GCCAGAGAGGACCTGAGAAAAGCGACTATCACTAGTGAATTAAGTCCAGCTGGAGG 767
DB 61 RPROGLUGLUGLYNTHGLULYSERGLYSEURTHSERAPLYALALATHRGLYOI 81
QY 768 AGCCCCAGGACGTGGAATAATATCTGCCGAGGACCAACTATGAGCAGAACTCTAC 827
DB 81 YALAPROGLYSERGLYASNSERLEUPROGLUGLNGLTHRWELUARGLINSERTH 101
QY 828 GAGCACAATCAACTGTGCTGTAGTGTCTTCTCCACAAAACCTCAGTGAACACTCGAATG 887
DB 101 RSEHTHRLAASNSERCYLEUVALALASERPROGLINASNSERSEGLUHSERANSE 121
QY 888 CAGCAATGTCGTGACAGTGGGCTTACTCAGATGTGAGATTGAGCAGCAATCGAA 947
DB 121 RSEHTHRLAASNSERGLYALATHRVALGILYEURTHGLINMETVALAPCYASPGILNGLSERY 141
QY 948 GAAAAACAATGTAGTGTGAGAGCAGGAGACCTGTGTAAGTCCCTCAAAATGGTTTA 1007
DB 141 ELYSASNSLYSCYSESERVALLYEASPLUGLULIA----- 152
QY 1008 AAATTTTAAATGATATGATGCGCTTACCTTTGTTAATCAATTAATGTTTGTGCTA 1067
DB 152 ----- 152
QY 1068 TACTTAACAATTTAGTTTAACTTGTAACCTTGACTAAACTCGGAGAGCTCGATCAA 1127
DB 152 ----- 152
QY 1128 AACAGACATTTTCTTGGAACCGTAATTAAGTCAATAAATAATTAATTCATCTTGATGA 1187
DB 152 ----- 152
QY 1188 ATGCATATCATGATGTACTCAAAACATCTCAAGAAAGACCTCAATTTGGATCACTAATT 1247
```


Db	152	-----	152	Db	221	-----	221
Qy	1248	AGTTGAGAAAAAATGCTGACTTTAAGAAATATTAATTTAAAAATTTGCTGAGTGA	1307	Qy	2328	GATTTACGAGTTATTTAATTTTTTAATGCTGATTTTAATAAATATCTTATTTATTCATTT	2387
Db	152	-----	152	Db	221	-----	221
Qy	1308	AATGATATATATGTCACAATATTTTATGTTAAAGCTAAAGCATTGTAAGCGGTG	1367	Qy	2388	TACATGATTAAAAATTGAAAAAATTTCAACGACAGTACACGGAAGAAAAAGATGCTCC	2447
Db	152	-----	152	Db	222	-Pro--ValLysIleGluArgIleSerAsnAspSerThrThiGluLysIleGlySerSer	240
Qy	1368	CTACGAGATGCTACTAGACGCGGTGAAAAAGCTAAATTTTATTTAAAAAGCTGCTAAT	1427	Qy	2448	TTGACAAATGAATATATGACAAATAGCATGAAAGCTGCATCATGTTGAATCCCATTTT	2507
Db	152	-----	152	Db	241	LeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGluLeuAsnProAspPhe	260
Qy	1428	ATTCCATTAACCATTAATGTCCTCATTTTCAGAAATTAAGTTCTTAATTAAGCAAAAGTCAAGC	1487	Qy	2508	ATCAATGAATCTTTAAATATATCCTGCAATTTTGAGATATTTGTAACGGAGTAGGACCA	2567
Db	153	-----	153	Db	261	IleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyValGlyPro	280
Qy	1488	AGCTGTGGCGGCTGGAAAAAGGTTCTATCATCCAGTTTGACTGTCAGGAGAAAGCCAC	1547	Qy	2568	ATACCCGAATGCGAGTTGGAGCGGGACGGGAAATTTTATTTGACTGCCAACGCCAATGGA	2627
Db	163	aaIaGlyGlyGlyCysGluThrGlySerThrSerLeuThrValLysGluLupProth	183	Db	281	IleProGlyIleGlyValIleGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAlaAsnGly	300
Qy	1548	CGATGCTTAGGCAAGTTTAGTAATATGAAAAAGAAAGAAAGAAATCATTCGCCAAC	1607	Qy	2628	ATCTCCTCGGGTAGCAATATTGTTGATTCATGCAACAGCAAAATCATATTTCTGTG	2687
Db	183	rhspValIleuGlySerLeuValAsnMetCysLysGluGluArgGluAsnHisSerProth	203	Db	301	IleSerSerGlySerSerAsnCysLeuAspTyrMetGluGluGlnGlnAsnHisIlePheVal	320
Qy	1608	GATGTCCTGTTGGTTGGTTTGCATTTGGTAAATGCAAGACAACTCGCTACACCGGG	1667	Qy	2688	TTTTCAACTCAGCTGGCCCAAAAGGGCGCGAATTCAGTTTAAAGCGTCAATTTCAACT	2747
Db	203	rmelSerProvalGlyPheGlySerIleGlyAsnAlaGlnAspAsnSerAlaThr-----	221	Db	321	PheSerThrGluLeuAlaAsnLysGlyAlaGluSerValLeuSerLysGlnPheGlnThr	340
Qy	1668	TAAGTTTAAAGATTCATATTAAGCAAAATTAACAAGATTAATGTCAGTTACCAATTTTA	1727	Qy	2748	ATTATTCGGTATCATCTGCACTCAGCTGCTACCAAAAAGCTTCCTGGAAGACTTTTATG	2807
Db	221	-----	221	Db	341	IleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeuGluAspPhePheMet	360
Qy	1728	TTTGATAGTCAAGAACTACTACTATAGCGATATCTCTGCTTTAATTTTATTTAATTAG	1787	Qy	2808	AAAAACCTTTAAAGATTAACAAGTTACAGCGGCACAAATTCCTGCTATGTCAGATA	2867
Db	221	-----	221	Db	361	LysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerValGlyMetProTyrIle	380
Qy	1788	GAAATACGATATTTCTAATTTGTAAATTAATAATTAATTAATTAATTAATTAATAAAC	1847	Qy	2868	GGCATGGGCGAGTTGGACTTAATCTCTCTAATTCCTGTAGCCCAATTAACAAACGACG	2927
Db	221	-----	221	Db	381	GlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLysIleThrGlnGlnGln	400
Qy	1848	CTTTTGATTAAGACATACCTCTCCAAAAATCAATATCATTTGGGAAACGAGTGTGTC	1907	Qy	2928	CCACATACAAAGACGCTAGCGCTATGAAACCCCAATTCATCAATGAAACAGACAA	2987
Db	221	-----	221	Db	401	ProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGlnHisGluAsnSerLys	420
Qy	1908	CCGAGGAGACTCTATTAACCTTTTGAGCTATCTGACTGACGCTACTAATAAATGA	1967	Qy	2988	CGTAGTACTGTAAACGCGCTTAGCAACTCTTTTGTGACCAAGCTGATCTATGGCAAC	3047
Db	221	-----	221	Db	421	ArgSerThrValSerAlaProSerAsnSerPheValAspGlnSerAspPrometGlyAsn	440
Qy	1968	TTAGTTAGGAAAAATGGGTGTAAATTTGTAGAAAGTTTTCATTTTAGAGAAAAATGTGATT	2027	Qy	3048	GAAACTGAATTGATGTGCTGGGAAAGCGGATCTCAAAACCCAGTATGCTTGACAAAC	3107
Db	221	-----	221	Db	441	GluThrGluLeuMetCysTrpGluGlyGlySerSerAsnHisSerLysSerGlyGlnAsn	460
Qy	2028	ATTTTATTAACCCCTTCAAGCGGAACACTACATTTGTTCTACGATATTTTGGAAAAACAA	2087	Qy	3108	TCACGAAATCATGTAGACAGTATCAGTACATCCAGGAGTCAAGAGCAATTAAGATAC	3167
Db	221	-----	221	Db	461	SerArgAsnHisValAspSerIleSerThrSerSerGluSerGlnAlaIleLysIleLeu	480
Qy	2088	TGTTTAAGTTGAAAGTGCCTATTAACAGAAATTCACGTTTCAATATCAACAGGTT	2147	Qy	3168	GAAGCAGCTGCGCTGATTTTGGACAGGTCACAAAAGAAAGCAGATCCTGGCTGACAACT	3227
Db	221	-----	221	Db	481	GluAlaAlaGlyValAlaAspLeuGlyGlnValThrLysGlySerAspProGlyLeuThrThr	500
Qy	2148	TTTGATTTAATTTTGATTAATAGAGAAATATACACTCTGATTAATATGTTTAATTGCA	2207	Qy	3228	GAAGCAACATTTGTATCATCTGCAAGAGATTAAAGTTTCAGAGCAAAACCTTAACACCA	3287
Db	221	-----	221	Db	501	GluAsnAsnIleValSerLeuGlnGlyValLysValProAspGluAsnLeuThrProGln	520
Qy	2208	TTTAAGTTCGACAAATCACAGACAGATTTTCATTTTGGGTATATATAGAAGTCCCTTC	2267	Qy	3288	CAGCGCAACATCGGAGAAAGACGTTGGCAAAATTAATAAATTAATGATCAATTTCTTTT	3347
Db	221	-----	221	Db	521	GlnArgGlnHisSerGluGluGluGlnLeuAlaLysIleLysLysMetLenglnPheLeuPhe	540
Qy	2268	ACACTTTTCTGGCGGCTTACACACTACGTAGAGTTCCGCCCGACGATGATTATATAGAT	2327	Qy	3348	CCTGAATAATGAGATTCAGTAGAGCTTAATGTAAGTTCACAGATTAACAAATTTCCAGA	3407
Db	221	-----	221	Db	541	ProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIleThrLysIleProGly	560

QY 3408 GATTTAATGATGGGATGTCGGGTGGCGAGCGGATCTATTATAATCCGATGCCA 3467
 |||||
 Db 561 AspLeuMetMetGlyMetSerGlyGlyGlyGlySerIleIleLeuProThrMetArg 580
 QY 3468 CAATCGCATATGCGCAGGTAAACGCAATGGAGCTCTTATCGCGCAAGTTGAGACTT 3527
 |||||
 Db 581 GlnLeuHleMetProIleAsnAlaLysSerGlnLeuLeuSerAlaThrSerSerGlyLeu 600
 QY 3528 TCGAAGATGTATGATCGAGGAGATGTTATGATGATATGGGATCCGTAATAGAGATG 3587
 |||||
 Db 601 SerGlnAspValMetHisProGlyAspValIleSerAspMetGlyAlaValIleGlyCys 620
 QY 3588 AATTAATATCAAAAAACAGTGTGCAATGTGATCTGAGTAGTGTTGCTGACTGGAGA 3647
 |||||
 Db 621 AsnAsnAsnGlnLeuThrSerValGlnCysGlySerGlyValGlyValValThrGlyThr 640
 QY 3648 ACTGCAAGCTGGAGTAAATGTCATATGATGCTCAAGCTCCGCGCCCGCAATGGCAT 3707
 |||||
 Db 641 ThrAlaAlaGlyValAlaAsnValAsnMetHisCysSerSerSerGlyValProAsnGlyAsn 660
 QY 3708 ATGATGGGAAGCTCTACGATATGCTAGCTCGTTGGCAACACAGCTGCAACGTCAATC 3767
 |||||
 Db 661 MetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThrSerCysAsnValIle 680
 QY 3768 GGAACGCGCCCGACATATGCTTAAGAAATTAAATCAAGATAGCCGAACCATTCACAT 3827
 |||||
 Db 681 GlyThrAlaProAspMetSerLysGlnValLeuAsnGlnAspSerArgThrHisSerHis 700
 QY 3828 CAAGGGGAGTCTCTCAATGAGTGTGCAAGATTCACATCAATTTTCCAGAAAGCC 3887
 |||||
 Db 701 GlnGlyGlyValAlaGlnMetGlnThrSerLysIleGlnHisGlnPheGlnGluAlaGly 720
 QY 3888 CTCAGAGGGGGCAAGCCGACAGACAAGTCACTGGAACGTAGTACCACAAACAGAAACCC 3947
 |||||
 Db 721 LeuLysGlyGlyLysProArgGlnValIleGlyThrValValProGlnGlnGlnThrPro 740
 QY 3948 TCTGATCTGGTGGAACTCGTTAAACAACAGAGTGCACCCCTGCAAGGTCCACCTCT 4007
 |||||
 Db 741 SerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeuGlnGlyProProPro 760
 QY 4008 CCTTACCATCTCCATCCAGAGATCTGGCTGATACCAATGACCACTCAATGCCCAATCC 4067
 |||||
 Db 761 ProIleHisSerIleGlnArgSerAlaSerValProIleAlaThrGlnSerProAsnPro 780
 QY 4068 TCGAGTCCAAACATGATCTCTCCGTCACCGGCAACGCGAGAGTCATGGGATGG 4127
 |||||
 Db 781 SerSerProAsnAsnLeuSerLeuProSerProArgThrThrAlaAlaValMetGlyLeu 800
 QY 4128 CCGACCAACTCTCTAGCATGATGAGAACAGATCATTTATCTGATCTGTTCCGCAAGCT 4187
 |||||
 Db 801 ProThrAsnSerProSerMetArgGlyThrGlySerLeuSerGlySerValProGlnAla 820
 QY 4188 AATACTTCCAGCGTTGACGGGACCAACAACAGTGTCTCAGCAACAAGACTGTTTT 4247
 |||||
 Db 821 AsnThrSerThrValGlnAlaGlyThrThrValLeuSerAlaAsnLysAsnCysPhe 840
 QY 4248 CAGGACAGACCCCATGCGCGTCAAAATCAAAATCGTAGTAAGTAATCCGATGTCAGC 4307
 |||||
 Db 841 GlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsnThrGlyLysSerSer 860
 QY 4308 GTTCTTAACGATTAATTAGCAGCAACCAAGTACCCTTATCTCATCTATCCCAAG 4367
 |||||
 Db 861 ValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSerHisLeuSerProLys 880
 QY 4368 GAATTTGAGTCTTTCCGTCAGTCTCTGCTGATATGTTATTTGTTAATTTTTTTAAA 4427
 |||||
 Db 881 GluPheGlySerPheGlyGlnSerSer----- 889
 QY 4428 GACAAATCAATATGAAATTCGTTAATTAATTAATATATATATATCAATCACTCGAAATTTG 4487
 |||||
 Db 889 ----- 889

QY 4488 ATAGAAAAATCAGGAATAGAAAAATAAATTATTTTCCGACCGCCCATCCATTTCTTG 4547
 |||||
 Db 889 ----- 889
 QY 4548 AATCCAAATTTTCGAGTGAATTTGTTAGAGTAATCTACTATTAAATTAACAGAAAAATT 4607
 |||||
 Db 889 ----- 889
 QY 4608 CATATCCGTTAATGAAAAATCACTATGTTTAAATAGAAATTTAAAAATATGTTTATATA 4667
 |||||
 Db 889 ----- 889
 QY 4668 ATATTTCTACAGGTATTAACATGAAAAAGTAGCGCACCAAGCCCAAGGCTCAGCGCTCAC 4727
 |||||
 Db 890 -----Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907
 QY 4728 CAGTAAATAGTCTATPAGCGCAATTAAGATGTAAGATTGCTGATCATCCAGTCCGTT 4787
 |||||
 Db 907 rovalAsnSerLeuIleGlnAlaAsnLysAspValArgPheAlaAlaSerSerProGlyP 927
 QY 4788 TTAAACCCGATCCACATATGCAAAAGCAATTCAAAATTGACATTTAAACGCTTAAATAG 4847
 |||||
 Db 927 heAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuAsnAlaTyLysMetG 947
 QY 4848 GCTCTACCAATATPACAGATGAGGTAAATATTTAATATTTTAACTTTTGTGTT 4907
 |||||
 Db 947 LysSerThrAsnIleGlnMetGlu--Arg----- 955
 QY 4908 AATTTATCTTCTTTTCAAGCTCAAGCATCAGCGCAAGGTGATCGGTCAATTTAGTGC 4967
 |||||
 Db 956 -----GlnAlaSerAlaGlnGlyGlySerValGlnPheSerArg 968
 QY 4968 GCGCTCCGATTAATATTCGCTAAATCCCAATAGTGGCAATCGCGCCCAACAACAAAGT 5027
 |||||
 Db 968 garGerSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgProProProAsnLysPhe 988
 QY 5028 GAACCAAACTTCGATCCATCTCTTCTTTGGCACAATATGTCCAACAACATAACAGTTG 5087
 |||||
 Db 988 ThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGlnGlnLeuThrSerCys 1008
 QY 5088 CGGTCCGACATGAGGTAGTCCAGCGGAACTGCTGTATGACATGATGGGGGGTCCGGG 5147
 |||||
 Db 1008 wAlaSerSerMetGlySerProIleGlyThrGlyGlyMetClnMetMetGlyGlyProGln 1028
 QY 5148 ACCGTCCGACATCAATATGACATGAAATATTTCCGGCATAGATGATCAGAAATAGA 5207
 |||||
 Db 1028 yProSerAspIleAsnIleGlnHisGlyIleIleSerGlyLeuAspGlySerGlyIleAsn 1048
 QY 5208 TACCATTAATCAAAATTAAGTGTCAATGAAATGTCGTAATGAATCAATAGGTCCTCCG 5267
 |||||
 Db 1048 pThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAsnSerMetGlyProArg 1068
 QY 5268 AATGCTGAATCTTAAATGCTGGTAGCAGCGGTCCAAATGACCGCTGAGCTTTAATCC 5327
 |||||
 Db 1068 gMetLeuAsnProLysPheMetCysValAlaGlyLysProAsnGlyProProGlyPheAsnArg 1088
 QY 5328 TAATTCGCCCAATGCTGATTAAGAGAGATTCATAGGGTCTGGCTGTGCTCAGCAAA 5387
 |||||
 Db 1088 oAsnSerProAsnGlyGlyLeuArgGlnAsnSerIleGlySerSerGlyCysGlySerAlaAs 1108
 QY 5388 CTCTTCAAACTTTCAAAGGGGTGTTCCACTGGTGCAGAAATGATGGGTGCAATGGCAAT 5447
 |||||
 Db 1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetGlyArgMetProVal 1128
 QY 5448 CAATTTGGTGTGAATTTCAATCCGAATTTGAGGTAAGAGGAGATACCCCAACACAT 5507
 |||||
 Db 1128 lAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSerThrProAsnThrIle 1148
 QY 5508 ACAATACATGCAAGTAAGGGCAGAGAACCCCAACAACATTAACAACATGAGCTAATA 5567
 |||||
 Db 1148 eGlnTyMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnAlaAsnGlyAlaAsnAs 1168
 QY 5568 TGTGCAATGCAACTGATGCTGGAATTTTGTGACAGAGTACGCTAACCTTCAAAATGGGTGC 5627

```

Db      1168 nValArgMetProSerLeuGlnPheLeuGlnArgTyrAlaAsnProGlnMetGlyAl
Qy      5628 TGTAGGCAATGGGTCCGCATATATGCCCATCATCAGCCAGGAGCGTACTCTCGAATGCC
Db      1188 aValGlyAsnGlySerProIleCysProProSerAlaSerAspGlyThrProGlyMetPr
Qy      5688 AGATATGATGGCGGAGCAGAGCCGAGGATGATGCTAATTAATCTTCCGAGAGCAACA
Db      1208 oGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSerSerGlyGlnGlnH
Qy      5748 CAGAAACAGATCAACAACAATCTCGGAGCAAGCAATGATTAATCTTCTTCAGAAATG
Db      1228 sGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAsnPhetheGlnAsnCy
Qy      5808 CAATCAATGCTCTATTGTTGACAGAGGGGTGATTAACCGGCGCATACGATCAATGAA
Db      1248 sAsnGlnMetSerIleValAspGlyGlnGlyLeuProGlyHisAspGlySerMetAs
Qy      5868 TATTGTCACCATCTATGATTAAGGGGATGCGTCACATGCGCATGGGCAATGTAAT
Db      1268 nIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMetArgProAsnValMe
Qy      5928 GGGTGCCTGGATGCGCACCCGTTAAACAGGCAATTCAGTTTGCACAGTCATCGATGCTAT
Db      1288 tGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGlnSerSerAspGlyI
Qy      5988 TGACTGTGTGGGATCCGTCATCATTTTTCATTAACGCTTCTCGCAACGCGCTGGACC
Db      1308 eAspCysValaGlyAspProSerSerPhePheThrAsnAlaSerCysAsnSerAlaGlyPr
Qy      6048 ACACATGTTGGATCAGACACAACAGGCGCATCGCCTAACACACAACATTAAGAATC
Db      1328 oHisMetPheGlySerIleGlnGlnAlaAsnGlnProLysThrGlnHisIleLysAsnI
Qy      6108 ACCTAGTGAATGTGTCAAAAACCAATCGGAGCTTGCAGTGCACAAGGGAGATCCAACT
Db      1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnIlyGlnIleGlnLe
Qy      6168 GCATGGGCAAGACATCGCAGGGTCACTTTTAATGGAACCTTAATAATTAATTAAT
Db      1368 uHisGlyGlnGlyHisIleAlaGlnIlyGlnSerLeuIleGlyProThrAsnAsnLeuMe
Qy      6228 GTCACTGCGGAGAGTGTCACTGCTACTAAGCGTGTCTCGGATCAATTCGAGGTCC
Db      1388 tSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIleAsnPheValaGlyPr
Qy      6288 CTCTTCTACGAGCTGAGATGATGCCAGCAATATCATAGTTTTCAGAGCAGTTATATGC
Db      1408 oSerSerThrAspLeuLysIleTyrAlaGlnGlnIlyThrIleSerPheGlnGlnIleuTyrAl
Qy      6348 TACCAACACCAAGAGTCAACAACAACAGATATGACAGCAGCAGCAGCAACATGAT
Db      1428 aThrAsnThrArgSerGlnGlnGlnGlnIleHisMetHisGlnGlnIleHisGlnSerAsnMetI
Qy      6408 AACAATGCCCGCAATTTATCAACCAATCCAAGTTCTTTTGTCAACAAA 6456
Db      1448 eThrMetProAsnLeuSerProAsnProThrPhePheValaAsnLys 1464

RESULT 4
ID      ABB58779 standard; protein; 1429 AA.
AC      ABB58779;
DT      26-MAR-2002 (first entry)
XX      Drosophila melanogaster polypeptide SEQ ID NO 3129.
DE      Drosophila; developmental biology; cell signalling; insecticide;
KM      pharmaceutical.
XX      Drosophila melanogaster.
OS

```

```

XX      MO200171042-A2.
PN      27-SEP-2001.
XX      23-MAR-2001; 2001WO-US009231.
XX      23-MAR-2000; 2000US-0191637P.
PR      11-JUL-2000; 2000US-00614150.
XX      (PEKE ) PE CORP NY.
PA      Venter JC, Adams M, Li PWD, Myers EW;
PI      WPI; 2001-656860/75.
XX      N-PSDB; ABL02882.
DR      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
PT      interactions.
PS      Disclosure; SEQ ID NO 3129; 21bp + Sequence listing; English.
XX      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC      sequences (ABU1840-ABU16175) and the encoded proteins (ABB57737-
CC      ABB72072). The sequence data for this patent did not form part of the
CC      printed specification, but was obtained in electronic format directly
XX      from WIFO at ftp.wipo.int/pub/published_pct_sequences
XX      SQ      Sequence 1429 AA;

Alignment Scores:
Pred. No.:      0      Length:      1429
Score:          7040.00      Matches:      1425
Percent Similarity: 71.09%      Conservative: 1
Best Local Similarity: 71.04%      Mismatches: 3
Query Match:     58.49%      Indels:      577
DB:              4      Gaps:      6

US-10-664-859-1 (1-6909) x ABB58779 (1-1429)
Qy      453 ATGCTCTGCAACAATGCCCCGAGTCCAAACCAACAGCCGCAACCAATCTCGAT 512
Db      1 MetLeuSerThrThrMetProArgSerProThrGlnGlnIleProGlnProAsnSerAsp 20
Qy      513 GCCTCTCAACAAGTGCATTTGATCAATCTTGAAGCAGCATCGGAATGGGAGACTCG 572
Db      21 AlaSerSerThrSerAlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSer 40
Qy      573 GCGGCGAGCAGAGATTCTCGAGAGACCTTAATAGGAACCTTTTCTACTTTGTGCGCC 632
Db      41 AlaAlaSerArgSerSerProLysTrpLeuAsnSerGlnProPheSerThrLeu-Ser-- 59
Qy      633 GGTAGAATTGTATTGATTTCTTTGTGCGGAATTATACAATTTCTGTGTTTCGAGA 692
Db      60 -----ProAs 61
Qy      693 TCAATATTAATTTAGCCGCAAGAAAGCACTGAGAAAGGAGCTATCACTAGATTA 752
Db      61 pGlnIleLysLeuThrProGlyGlnIlyThrGlyLysSerGlyLeuSerThrSerAspLys 81
Qy      753 AGCTGCCACTGAGAGAGCCCAAGGAGTGAATAATATCTGCCGAGGAGCAAACTATGCT 812
Db      81 sAlaAlaThrGlyGlyAlaProGlySerGlyAsnAsnLeuProGlyGlnIlyGlnThrMetLe 101
Qy      813 AAGGCAAACTTACAGACCAATCAACTCTGCTTACAGTGTCTTTCACAAAATCTCAG 872
Db      101 uArgGlnAsnSerThrSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSe 121

```

QY 873 TGAACACTGGAAATAGCAGCAATGTGTCTGCTACAGTGGCCTTACTCAGATGGTAGATGG 932
| | | | |
Db 121 rgluh1SerAenSerSerAenValSerIalThrValGlyLeuThrGluMetAlaIlePcy 141
| | | | |
QY 933 TGAACGCAATCGAATAAAAAACAATGTAGTGTGAAGACGAGGAAGCTGGTAAAGCTCG 992
| | | | |
Db 141 sApGluGlnSerLysValSerLysCysSerValLysApGluGlnAla----- 157
| | | | |
QY 993 CCTACAAATGTGTTAAAAATTTTAAATGTATGGCGTCACTTGTATCATTTAATT 1052
| | | | |
Db 157 ----- 157
| | | | |
QY 1053 GTTTTTTTTGTCTATACCTTACATTTTAGTTTAACTTGAACTGACTAAACCTCG 1112
| | | | |
Db 157 ----- 157
| | | | |
QY 1113 CGAAGCTCGGATCAAAAACAGACATTTTCTTGGAACCGTAATTAACTCATMAAAATATTA 1172
| | | | |
Db 157 ----- 157
| | | | |
QY 1173 ATTCACTTGATGAAATGATCATAGATGTACTCAAAACATCTCAAGAAAGACTCAAA 1232
| | | | |
Db 157 ----- 157
| | | | |
QY 1233 TTGGATCAACTAATTAGTTTGAAGAAAAATGCTGTACTTTAAGAATATATTAAATTAA 1292
| | | | |
Db 157 ----- 157
| | | | |
QY 1293 AAATTTGCTGAGTAATGATATATATATAGTCACAAATATTTTAGTTAACTGCTAAAGCA 1352
| | | | |
Db 157 ----- 157
| | | | |
QY 1353 TTTTGATAGCGTGTCTAGCAGATGCTACTAGACGGGTAAAGTAATTTTATTTT 1412
| | | | |
Db 157 ----- 157
| | | | |
QY 1413 AAAAGTGTCTAATATTCATATCCATTAATGTCCCATTTGAGAAATAAGTTCTAATA 1472
| | | | |
Db 158 ----- 163
| | | | |
QY 1473 AGCAAAAGTCAAGCAGCTGTGGCGGCTGCCAAACAGGTTCTACATCCAGTTTGACTGT 1532
| | | | |
Db 163 sAlaLysGlyGlnAlaIleGlyGlyCysGluThrGlySerThrSerLeuThrVal 183
| | | | |
QY 1533 CAAGAAGAACCCACCGATGTCTTAGCAGAGTTAGTAAATATGAAAAAGAAAGAGAGA 1592
| | | | |
Db 183 llySGluGluProThrAspValIleuGlySerLeuValAsnMetLysGluIleArgG 203
| | | | |
QY 1593 AAATCAATTGCGCAACGATGTCCCTGTGTGTTTGGTTCAATTGGTAATGCACAGAGCAA 1652
| | | | |
Db 203 uAenH1SerProThrMetSerProValGlyPheGlySerIleGlyAsnAlaIleAspAs 223
| | | | |
QY 1653 CTCGGCTACCCGGGTAAAGTTTAAAGATCCATATTAAGCAATAACAAGAAATTATGT 1712
| | | | |
Db 223 nSerIalThr----- 226
| | | | |
QY 1713 CAGTTACCAATTTTATTGTAGTCAAAAGACTACTATAGCGATATCTCGTGCCTTTAA 1772
| | | | |
Db 226 ----- 226
| | | | |
QY 1773 TTTTATTTTATAGAAATAGCAATATTTCTAATTTGTAAAAATAAAATTGATTAAATTAA 1832
| | | | |
Db 226 ----- 226
| | | | |
QY 1833 CTAGAATTTAAAAACCTTTTGAATTAGACATACCCCTTCAAAAATCAGTAATCATTTGGG 1892
| | | | |
Db 226 ----- 226
| | | | |
QY 1893 AACGAGATGTGTGCCGAGAGAGACTATATAAACTTTGAGCTATCTGATCTGCA 1952
| | | | |
Db 226 ----- 226

QY 1953 CGCTACTAAATAATGATTAGTTAGGAAAAATGGGTGAATTTTGTAGAAAGTTTCATTTT 2012
| | | | |
Db 226 ----- 226
| | | | |
QY 2013 AGAAGAAATGTATTTATTTTATTAACCCCTTCAGAGCGGAAGTACTATTTGTTCTACGATA 2072
| | | | |
Db 226 ----- 226
| | | | |
QY 2073 TTTTGGAAAAACAATGTGTTAAGTTGAAAGTGGCTATATAACAAGATTTCCAGGTTTCA 2132
| | | | |
Db 226 ----- 226
| | | | |
QY 2133 AATTAACACGAGTTTGTGATTAAATTTTGATTAATGAGAAATTAATCACACTTCAGTTA 2192
| | | | |
Db 226 ----- 226
| | | | |
QY 2193 AAATGTTTAATTCGATTAAAGTCGGCAATCAAGAGATTTCCATTTTGGGTGATAT 2252
| | | | |
Db 226 ----- 226
| | | | |
QY 2253 ATAGAAGTCGCTTGACACTCTTCTGGCGGCTTCACCACTAGTGAAGTTCGCCCGCA 2312
| | | | |
Db 226 ----- 226
| | | | |
QY 2313 GTGATTTATATAGATGATTACGAGTTAATTTTATATGTGTTATTTAATAATAT 2372
| | | | |
Db 226 ----- 226
| | | | |
QY 2373 CTTATTTATTCATTTTACATAGTTAAATTTGAAGAATTTCAAGACAGTACACGGAA 2432
| | | | |
Db 227 ----- 240
| | | | |
QY 2433 AAAAAAGATCGTCTTGAACAATGAAATATGACGAATAGAGATGGAAGCTGCAATCG 2492
| | | | |
Db 241 LysLysGlySerSerLeuThrMetLeuAsnAspGluMetSerMetGluGlyCysAsnGln 260
| | | | |
QY 2493 TTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCTTGCAATTTTCAGCATATTAGTA 2552
| | | | |
Db 261 LeuAsnProAspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuVal 280
| | | | |
QY 2553 AGCGGAGTAGGACCAATATCCGGAATCCGAGTTGGAGCGGGGACGGAAATTTATGACT 2612
| | | | |
Db 281 SerGlyValGlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThr 300
| | | | |
QY 2613 GCCAAGCCCAATGGAATCTCTCGGATGACAGTAATTTGTTGATTTACATGCAACAGCAA 2672
| | | | |
Db 301 AlaAsnAlaAsnGlyIleSerSerGlySerSerAsnCysLeuAspIleMetGlnGlnGln 320
| | | | |
QY 2673 AATCAATATTTGTTGTTTCAACTCAGCTGGCCAAACAAAGGGCCGAATCAAGTTTAAAC 2732
| | | | |
Db 321 AsnH1sIlePheValPheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValIleuSer 340
| | | | |
QY 2733 GGTCAATTTCAACTATTTATGCGTATCAGTCACTGACCTCAGCTGTCAAAAAAGCTTCG 2792
| | | | |
Db 341 GlyLysPheGlnThrIleIleAlaIleThrIleCysThrGlnProAlaThrLysSerPheLeu 360
| | | | |
QY 2793 GAAGACTTTTATGAAAAACCTTTAAAGATTAAACAAGTTACAGCGGACAAATTCGCT 2852
| | | | |
Db 361 GluAspPhePheMetLysAsnProLeuLysIleAsnLysLeuGlnArgH1sAsnSerVal 380
| | | | |
QY 2853 GGTATGCCATGGATAGGCAATGGGGCAGGTTGGACTCACTCTCTTAATCTCTGAGCCAAA 2912
| | | | |
Db 381 GlyMetProIleIleGlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLys 400
| | | | |
QY 2913 ATTAACAACAGACAGCACTACAAGACCGTAGGCTTATTAAGAACCCCAATTCATTCAA 2972
| | | | |
Db 401 lIeHnGlnGlnGlnProH1sThrLysThrValGlyLeuLeuLysProGlnPheAsnGln 420
| | | | |
QY 2973 CATGAACACGACAAAGTAGTACTGTAAAGCGGCGCTAGCAACTCTTTTGTGACCAAGTCT 3032
| | | | |
Db 421 H1sGluAsnSerLysValSerThrValSerAlaProSerAsnSerPheValAspGlnSer 440
| | | | |
QY 3033 GATCTATGGGCAACGAACGAAATTGATGTGCTGGGAAGGGGAGTCTCCAAACACCAAGT 3092

Db 441 AspProMetGlyLeuAngIuThrGluLeuMetCysEtrpGluGlySerSerAenThSer 460
QY 3093 AGGTCTGGACAAAACTCAGCAAAATCATGTGACAGTATCATGTACATCCAGGACTCAG 3152
Db 461 ArgSerGlyGlnAsnSerArgAsnH1sValAspSerL1eSerThrSerSerGluSerGln 480
QY 3153 GCATATAAGTACTGGAGAGCGTGGCTGATTGGGACAGGCACAAAAAGGAAGCGAT 3212
Db 481 AleIleYsIleLeuGluAlaIaGlyValAspLeuGlyGlnValThrLySgLySerAsp 500
QY 3213 CCTGGCTGCAACATGAAAAACAATTGTATCATCTGCAAGAGAGTTAAAGTTCCAGACAA 3272
Db 501 ProGlyLeuThrThrGluAsnAnIleValSerLeuInGlyValLySValProAspGlu 520
QY 3273 AACCTTACACCAACACGCGGCAACATCGGAAAGACAGTGGCAAAAATAAAAAATG 3332
Db 521 AsnLeuThrProGlnGlnArgGlnH1sArgGluGluGlnLeuAlaYsIleLySgYsMet 540
QY 3333 AATCAATTTCTTTTCCCTGAAAAATGAGATTCAGTAGAGCTAATGTAACTCACAATA 3392
Db 541 AsnGlnPheLeuPheProGluAsnGluAsnSerValG1yAlaAsnValSerSerGlnIle 560
QY 3393 ACAAAAATTCAGAGAGATTTAATGATGGGATGTCCGGTGGCGAGCGGAGATCTATTATA 3452
Db 561 ThrLyIleIleProGlyAspLeuMetMetGlyMetSerGlyGlyGlySerIleIle 580
QY 3453 AATCCGACGATGGACACATGCTATGTCCAGGTACGCCAAATCGAGCTTATCCGGC 3512
Db 581 AsnProThrMetArgGlnLeuH1sMetProGlyAsnH1aLySerSerGluLeuLeuSerAla 600
QY 3513 ACAAGTTCAGAGACTTGGAGAGATGTAAATGATCATCCAGGGAGTGTATATAGATATAGGT 3572
Db 601 ThrSerSerGlyLeuSerGluAspValMetCh1ProGlyAspValIleSerAspMetGly 620
QY 3573 GCCGTATAGATGATATAATTAATCAAAAAACAGTGTGCAATGTGATCTGAGTAGT 3632
Db 621 AlaValIleGlyCysAsnAsnAnGlnLyThrSerAlaGlnCysGlySerGlyValGly 640
QY 3633 GTTGTCACTGGAACAATGCACTGAGTGAAGTAAATGTCAATATGCAATTGCTCAAGCTCCGGC 3692
Db 641 ValValThrGlyThrThrAlaAlaGlyValAsnValaAsnMetCh1sCysSerSerSerGly 660
QY 3693 GCCCGAATGGCAATATGATGGAGAGCTACAGATATGCTAGCTGCTGGTGGCAACA 3752
Db 661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
QY 3753 AGCTGCAACGTCAATCGGAACGCGCCAGATATGTCTAAGAGAGTTTAAATCAAGATAGC 3812
Db 681 SerCysAsnValIleGlyThrAlaProAspMetSerLySgIuValLeuAnGlnAspSer 700
QY 3813 CGAACCCATTCAATCAAGGGGAGATTGCTCAAAATGAGTGGTGAAGATTCAACATCA 3872
Db 701 ArgThrHisSerHisGlnGlyValaIaGlnMetC1uTrpSerLyS1leGlnH1sGln 720
QY 3873 TTTTGGAAAGAGCGCTCAAGGGGGGAGCCAGCCAGACAACTGGAACCTGAATCTAGTACA 3932
Db 721 PhePheGluGluTrpGlyLeuLySgLyLySgProArgGlnValThrGlyThrValaPro 740
QY 3933 CAACAGCAAAACCCCTTCTGATCTGTGTGAACCTCGTTAAACAACAGTGGCAACCCCTG 3992
Db 741 GlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAnGlnAnGlnValaArgProLeu 760
QY 3993 CAAGGTTCACCTCTCTTACCACTCCATCCAGAGATCTGCGTCAAGTACCAATAGCACT 4052
Db 761 GlnGlyProProProProLyThrHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
QY 4053 CAATCGGCAATCCCTGAGTCCAAACAATCTATCTTCCCGTCAACGCGGAGCAACCGCA 4112
Db 781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThrAla 800
QY 4113 GCAGTCATGGGATGGCCGACCAACTCTCTAGCATGATGAACAGGATCATTAATCTGGA 4172

Db 801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
QY 4173 TCTGTCCGCAACTTAATATTGGACGGTTCCAGGAGCGCAACAACAGTCTCTCAGCA 4232
Db 821 SerValProGlnAlaAsnThrSerThrValaGlnAlaGlyThrThrValaLeuSerAla 840
QY 4233 AACAGAACTGTTTTCAGGAGACACCCCATCGCCCTCAATCAAAATCGTAGTAGAAAT 4292
Db 841 AsnLyAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAnAspSerArgAsn 860
QY 4293 ACCGATCGTCAAGCGTCTTACGCAATACCTAAGACAGCAACCAAGTACCCCTTATCT 4352
Db 861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
QY 4353 CATCTATCCCAAGAAATTTGAGTCTTTCGTCAGTCCCTGCTGCTGTATATATTG 4412
Db 881 HisLeuSerProLySgIuPheGlnSerPheGlyGlnSerSer----- 894
QY 4413 TTTAATTTTAAAGCAAAATCAATCAATATGCAATTTGCGTTAATTAATTAATTAACA 4472
Db 894 ----- 894
QY 4473 TAACTGGAAATTTGATAGAAAAATCAGAAATAGAAAAAATAATTAATTTCCGACCG 4532
Db 894 ----- 894
QY 4533 CCCATCCATTTCTTGAATCCAAATTTCTGAGTAGATTGTTAGATATATCTACTATTAAA 4592
Db 894 ----- 894
QY 4593 TTAACACAGAAATTCATATCCGTTAATGAAAAATCATATTGTTTAATAGAAATTTAAA 4652
Db 894 ----- 894
QY 4653 AATATGTTATTATTAATTAATTTCTACAGGTGATTAACATGAAGATGAGCGCAACCCAC 4712
Db 895 -----Ala--GlyAspAsnMetLySgSerArgArgProSerProG 907
QY 4713 AGGTGACGCGGTACACAGTAAATAGTCTAATAAGCAAAATAAGATGTACGATTGCTG 4772
Db 907 InGlyGlnArgSerProValaAsnSerLeuIleGluAlaAsnLyAspValaArgPheAla 927
QY 4773 CATCGAGTCTGCTTTTAAACCCGATCCACATATGCAATGCAAAAGCAATTCATAGCATTAA 4832
Db 927 IAserSerProGlyPheAsnProHisProHisMetGlnSerAsnSerAlaSerAla 947
QY 4833 AGCCTATATAAATGGGCTCTACCAATATACAGATGAGAGTAAATTAATTAATTTATT 4892
Db 947 snIaLyLySgMetGlySerThrAsnIleGlnMetGlu--Arg----- 960
QY 4893 TAAAGTTTGTGTAAATTAATCTCTTTTTCAGCGTCAAGCATACGCGGAAGTGTGATC 4952
Db 961 -----GlnAlaSerAlaGlnGlyLyse 968
QY 4953 CGTACAAATTAATGTCGCGGCTCCGATATAATTCGCTAAATCCCAATAGTGGCAATCGGCC 5012
Db 968 rValGlnPheSerArgArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgPr 988
QY 5013 GCCACCAACACAGATACCCAAAATCTTGATCCAAATCTCTTTGGCACAAAATGTCACA 5072
Db 988 oProProAsnLyMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerG1 1008
QY 5073 ACAACTTAACAAGTTGGGTGTCGACGATGCGGTATCCAGCGGAACTGGTGTATGACAT 5132
Db 1008 ngInLeuThrSerCysValaSerSerMetGlySerProAlaGlyThrGlyGlyMetCh1rMe 1028
QY 5133 GATGGGAGGTCGCGGACCGTCCGACATCAATATTTGAGCATGAGTAATTTCCGAGACTAGA 5192
Db 1028 tMetGlyGlyProGlyProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAs 1048
QY 5193 TGGATCAAGAAATGATACCAATTAATCAAAATTAATCTGATTCATGATGATGCTTAATGAA 5252
Db 1048 pGlySerGlyIleAspThrIleAsnGlnAsnAnCysHisSerMetAsnValaValMetAs 1068

```

QY 5253 CTCATGGGTCCCGCATGCTGATCTTAAATGTGCTAGCAGGCGGTCCAATGGACC 5312
    |||||
Db 1068 nserMetGlyProAArgMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyLyr 1088
QY 5313 GCGTGGCTTTAATCTTAATTTCCCGCATGGTGGATTAAAGAGAAATTCATAGGGTCTTG 5372
    |||||
Db 1088 oProGlyPheAsnProAsnSerProAsnGlyLeuAArgGluAsnSerIleGlySerGly 1108
QY 5373 CTGTGGCTCAGCAAACTTTTCAAACTTTTCAAGGGGTTGTTCCAGCTGGGCCGAAATGAT 5432
    |||||
Db 1108 yCysGlySerAlaAsnSerSerAsnPhenGlnGlyValProProGlyAlaAArgMetHe 1128
QY 5433 GGGTCGAATGCCAGTCAATTTTGGTTGCAATTTCCAAATTCGAAATATTCAGGTAAAGCGAG 5492
    |||||
Db 1128 tGlyAArgMetProValAsnPhenGlySerAsnPhenProAsnIle----- 1143
QY 5493 TACCCCAAAACCATATCATTCATGCCAGTAAGGAGCAGAACCCAAACATATACAA 5552
    |||||
Db 1143 ----- 1143
QY 5553 CAATGAGCTAATATATGCGAAATGCCACCTAGTCTGAAATTTTGCAGAGGACGCTAA 5612
    |||||
Db 1144 -----GlnArgTyrAlaAs 1148
QY 5613 CCTCAATGGGTGCTGTAGGCAATGGGTGCGCAATATGCCCAATCAGCCAGCGAGCG 5672
    |||||
Db 1148 nProGlnMetGlyAlaValAlaGlyAsnGlySerProIleCysProProSerAlaSerAspGly 1168
QY 5673 TACTCTGGAATGCCAGGATTTATGCGCGGAGCAGAGAGCGAGATGCTAATGAATTC 5732
    |||||
Db 1168 yThrProGlyMetProGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetLeuAsnHe 1188
QY 5733 TTCGCGAGAGCAACACCAGAACAGATCAAAACAATCTCGGGGCAAGCAATGATTTAA 5792
    |||||
Db 1188 tSerGlyGlyGlnGlnIleAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAs 1208
QY 5793 CTTCTTTGCAATTTGCAATCAATGTCCTATTGTTGAAGAGAGGTGATTTACCCGCGCA 5852
    |||||
Db 1208 nPhePheGlnAsnGlyAsnGlnMetSerIleValaArgGlyGlyGlyLeuProGlyIle 1228
QY 5853 TGACCGATCAATGAATATTGTTGTCACACCTATATGATTAAGGGGATCCGTTCCATCCAT 5912
    |||||
Db 1228 sAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgProIleAlaMe 1248
QY 5913 GCGGCCAATGTATGGGTGCGCGAGTCCACCGCTTAAACAGCAATTCAGTTTGACAA 5972
    |||||
Db 1248 tArgProAsnValMetGlyAlaAArgMetProProValAsnArgGlnIleGlnPheAlaGly 1268
QY 5973 GTCATCGAATGATTAATGACTGTGTGCGGAGTCCGTATCATTTTTCATTAACGCTTCTTG 6032
    |||||
Db 1268 nSerSerAspGlyIleAspCysValAlaGlyAspProSerSerPhePheThrAsnAlaSerC 1288
QY 6033 CAACACCGCTGACCAACACATGTTTGGATTCAGACAAACAGGCCAATCAGCTTAAGACAA 6092
    |||||
Db 1288 sAsnSerAlaGlyProIleMetHeGlySerAlaGlnGlnAlaAsnGlnProLysThrGly 1308
QY 6093 ACACATTAAGAATACATACCTAGTGAATGTGTCAAAACCAATCGGAGCTTGCACTGGACAA 6152
    |||||
Db 1308 nHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValaAlaGly 1328
QY 6153 AGGGCAGATTCAACTGCATGGGCAAGACATGCCAGGGTCACTTTTAATTTGACCTTAC 6212
    |||||
Db 1328 nGlyGlnIleGlnLeuIleGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProTh 1348
QY 6213 TATATTAATTTAATATGCACTGCCGAGTGCATGCTACATCAATTAAGGTCTCTGGCAT 6272
    |||||
Db 1348 tAsnAsnAsnLeuMetSerThrAlaGlySerValSerAlaThrAsnGlyAlaSerGlyIle 1368
QY 6273 CAATTCGTAGAGTCCCTCTTCTACGAGCTGAAGTAAAGCCACATATCATATGATTTTCA 6332
    |||||
Db 1368 eAsnPhenValGlyProSerSerThrAspLeuLysTyrAlaGlnGlnIleThyHisSerPheGly 1388

```

```

QY 6333 GCAGCATTTATATGTCTACCAACACAGAGATGCAACAAACAGATATGACACAGACA 6392
    |||||
Db 1388 nGlnGlnLeuTyrAlaThrAsnThrAArgSerGlnGlnGlnIleHisMetHisGlnGlnIle 1408
QY 6393 CCAGAGCAACATGATTAACAATGCCCGCGCAATTTATACCAATTCACAGCTTCTTTGTCAA 6452
    |||||
Db 1408 gGlnSerAsnMetIleThrMetProAsnLeuSerProAsnProThrPhePheVala 1428
QY 6453 CAAA 6456
    |||||
Db 1428 nLys 1429

RESULT 5
ABW01532
ID ABW01532 standard; peptide; 112 AA.
XX
AC ABW01532;
XX
DT 15-JUN-2004 (first entry)
XX
DE Drosophila species legless (lgs) peptide #5.
XX
KW legless protein; lgs; cell fate disorder; blood disease; gene therapy;
XX cancer; tissue regeneration; tissue repair; cytostatic.
XX
OS Drosophila sp.
XX
PN US2003114413-A1.
XX
PD 19-JUN-2003.
XX
PF 19-DEC-2002; 2002US-00322579.
XX
PR 28-JUL-2000; 2000US-0221502P.
XX
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2003-829432/77.
XX
PT Novel lgs polypeptide useful for isolation of lgs-binding proteins,
XX diagnosing disorders of cell fate, treating diseases such as cancer.
XX
PS Claim 28; Fig 7B; ODP; English.
XX
CC The invention relates to novel legless (lgs) proteins and polynucleotides
CC encoding such proteins. Lgs sequences are useful for the treatment of
CC disorders of cell fate such as differentiation or proliferation. The
CC invention is used to treat blood disease or a cancerous condition
CC characterized by over-stimulation of the Wnt pathway such as colon,
CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and
CC is administered to prevent progression from a pre-neoplastic or non-
CC malignant condition to a neoplastic or malignant state. It is
CC administered to promote tissue regeneration and repair. The invention is
CC also useful in the therapy of diseases cost by an over-activation of Wg
CC pathway. It is useful for reducing lgs gene expression in an invertebrate
CC or vertebrate organism or an invertebrate or vertebrate cell line. The
CC invention is also useful in gene therapy. The present sequence is
CC Drosophila species legless (lgs) peptide
XX
SQ Sequence 112 AA;
XX
Alignment Scores:
Pred. No.: 9 01e-40 Length: 112
Score: 570.00 Matches: 112
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.74% Indels: 0
DB: 7 Gaps: 0
US-10-664-859-1 (1-6909) x ABW01532 (1-112)

```

QY 4029 TCTGCGTACGATACCAATAGCCAGCTCAATCGCCCAATCGCTGAGTCCAAACATCTATCT 4088
 |||||
 Db 1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnLeuSer 20
 |||||
 QY 4089 CTCCTGTCACCGCGGAGCAACCGAGCATGAGGATGGCCGACCACTCTCTAGCATG 4148
 |||||
 Db 21 LeuProSerProAlaGlnThrAlaAlaValMetCglyLeuProThrAsnSerProSerMet 40
 |||||
 QY 4149 GATGGAACAGAGATCATTTATCTGATCTGTTCCGAGACTAATCTTGAAGGTTCAAGCA 4208
 |||||
 Db 41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60
 |||||
 QY 4209 GGACACCAACAGAGTCTCTGAGCAACAGAACTGTTTTCAGGAGCAGACCCCACTGCGCG 4268
 |||||
 Db 61 GlyThrThrThrValLeuSerAlaAsnLysAsnGlyPheGlnAlaAspThrProSerPro 80
 |||||
 QY 4269 TCAAAATCAAAATGCTAGTACGAAATACCGAGATCGTCAAGCGTTTTCAGCATTAATAGC 4328
 |||||
 Db 81 SerAsnGlnAsnAlaArgSerAlaGlnThrGlySerSerSerValLeuThrHisAsnLeuSer 100
 |||||
 QY 4329 AGCAACCCCAAGTACCCCTTATCTCATCTATCCCA 4364
 |||||
 Db 101 SerAsnProSerThrProLeuSerHisLeuSerPro 112
 |||||

RESULT 6
 ADJ71898
 ID ADJ71898 standard; protein; 112 AA.
 AC ADJ71898;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Fruit fly legless (lgs) peptide fragment #5.
 XX
 KM Fruit fly; legless; lgs; cell differentiation disorder;
 KM cell proliferation disorder; cancer; wnt pathway; medulloblastoma; colon;
 KM breast; head; neck; brain; thyroid; skin; blood disease;
 KM tissue regeneration; tissue repair; cytostatic.
 XX
 OS Drosophila melanogaster.
 OS
 XX
 PN US2004038901-A1.
 XX
 PD 26-FEB-2004.
 XX
 PF 22-SEP-2003; 2003US-00664859.
 XX
 PR 28-JUL-2000; 2000US-0221502P.
 PR 27-JUL-2001; 2001US-00915543.
 XX
 PA (UYZU-) UNIV ZURICH.
 XX
 PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
 XX WPI; 2004-203288/19.
 XX
 PT Novel polypeptide sharing one or more homologue amino acid domains with
 PT Legless protein being functional homologue of Legless, useful for
 PT diagnosing disorders of cell fate.
 XX
 PS Disclosure; SEQ ID NO 10; 62pp; English.

CC are useful for treating disorders of cell fate, which involves
 CC administering therapeutic compounds chosen from invertebrate and
 CC vertebrate lgs protein homologues or fragments, antibodies, antibody
 CC fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,
 CC small peptides or chemical and natural compounds being capable of
 CC interfering with lgs function, synthesis and degradation. The disorders
 CC are related to cell differentiation or cell proliferation. The compound
 CC is administered to treat a cancerous condition by preventing progression
 CC from a pre-neoplastic or non-malignant condition to a neoplastic or
 CC malignant state. The cancerous condition is characterized by over-
 CC stimulation of the Wnt pathway and is medulloblastoma or cancer of the
 CC colon, breast, head and neck, brain, thyroid or skin. The therapeutic
 CC compound may also be administered to a blood disease to promote tissue
 CC regeneration and repair. This sequence represents a Drosophila lgs
 CC peptide fragment of the invention.

XX
 XX SQ Sequence 112 AA;
 XX
 Alignment Scores:
 Pred. No.: 9,01e-40 Length: 112
 Score: 570.00 Matches: 112
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.74% Indels: 0
 DB: 8 Gaps: 0

US-10-664-859-1 (1-6909) x ADJ71898 (1-112)

QY 4029 TCTGCGTACGATACCAATAGCCAGCTCAATCGCCCAATCGCTGAGTCCAAACATCTATCT 4088
 |||||
 Db 1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnLeuSer 20
 |||||
 QY 4089 CTCCTGTCACCGCGGAGCAACCGAGCATGAGGATGGCCGACCACTCTCTAGCATG 4148
 |||||
 Db 21 LeuProSerProAlaGlnThrAlaAlaValMetCglyLeuProThrAsnSerProSerMet 40
 |||||
 QY 4149 GATGGAACAGAGATCATTTATCTGATCTGTTCCGAGACTAATCTTGAAGGTTCAAGCA 4208
 |||||
 Db 41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60
 |||||
 QY 4209 GGACACCAACAGAGTCTCTGAGCAACAGAACTGTTTTCAGGAGCAGACCCCACTGCGCG 4268
 |||||
 Db 61 GlyThrThrThrValLeuSerAlaAsnLysAsnGlyPheGlnAlaAspThrProSerPro 80
 |||||
 QY 4269 TCAAAATCAAAATGCTAGTACGAAATACCGAGATCGTCAAGCGTTTTCAGCATTAATAGC 4328
 |||||
 Db 81 SerAsnGlnAsnAlaArgSerAlaGlnThrGlySerSerSerValLeuThrHisAsnLeuSer 100
 |||||
 QY 4329 AGCAACCCCAAGTACCCCTTATCTCATCTATCCCA 4364
 |||||
 Db 101 SerAsnProSerThrProLeuSerHisLeuSerPro 112
 |||||

RESULT 7
 ABB11808
 ID ABB11808 standard; peptide; 1435 AA.
 AC ABB11808;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human BCL9 homologue, SEQ ID NO:2178.
 XX
 KM Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KM haematopoietic regulation; tissue growth; immunomodulator; activin;
 KM inhibin; chemokinesis; chemokinesis; thrombolytic; oncogenesis;
 KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KM chronic inflammatory condition; proliferative retinopathy;
 KM atherosclerosis; coronary heart disease; arterial ischemia;
 KM bone disorder; osteoporosis; vascular growth disorder;
 KM tissue regeneration; wound healing; infection; immune disorder;
 KM cell culture; drug screening; gene therapy; antiinflammatory;
 KM antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

QY	4958	AAATTAGCGGC-----GCTCCGAATAATTCGGTAAATCCGATA	4999
Db	931	GlnleuAlaGlyMetleuAlaGlyProAlaAlaAlaAlaSerIleuYSerProProAl	950
QY	5000	GTGGCAATCGGCGCC-----	5015
Db	951	LeuGly-SerAlaAlaAlaSerProValIleuYSerProSerIleuProAlaProSe	970
QY	5016	-----ACCAACAAGATGACCCCAAACTCGATCCCAATCTCTTTGGCACAA	5065
Db	970	rProGlyTrpThrSerSerProGluProIleuIndinSer-----ProG1	985
QY	5066	TGTCCCAACAACTAACAAGATTGCGGTCCGATCGATGATGCCAGCCGAATGGTGTA	5125
Db	985	YIleProProAsnHisIleYbAlaProIleu-ThrMetAlaSerProAlaMetIleuIleuYanv	1005
QY	5126	TGACGATGATGGGGGCTCG-----GGACGCTCGACATCAATATTGAGC	5170
Db	1005	AlGluSerGlyYGIProProProProThralaSerGlnProIleAlaSerValIleProG	1025
QY	5171	ATGGAATATTTTCGGGACTAGATGGATCA-----GGAATGATATCCATAATCAATA	5224
Db	1025	lySerIleuProSerSerThrProTyTrpMetProProGluProThrIleuSerGlnIleu	1045
QY	5225	ACTGTCATTCATGAATGATCGTAAATGAATCAATGGGTCGCCGATGTCGATCT	5280
Db	1045	roLeuSerIleMetMetSerArgMetSerIleYpHeAlaMetProSerSerIleuProGlyT	1065
QY	5281	-----AAATGTGCGTAGCAGAGCGGGCTCCAAATGACACCGCTGGCTTTA	5323
Db	1065	yrAsnHisAspAlaIleIleYThrValAlaSerSerAspAspAspSerProAlaArgS	1085
QY	5324	ATCCTAATTCCTCCCAATGGTGGATTAAAGAGAATTCATAGGGCTGCTGTGGCTCAG	5383
Db	1085	erProAsnIleuProSer-----MetAsnAsnMetProGlyMetCylIleAsnThrG	1102
QY	5384	CAAA-----	5388
Db	1102	IAsnProArgIleSerGlyProAsnProValAlaProMetProThrIleuSerProMetG	1122
QY	5389	-----TCTTCAACTTTCAGGGCTT	5409
Db	1122	lyMetThrGlnProIleuSerHisSerAsnGlnMetProSerProAsnAlaValGlyProA	1142
QY	5410	-----GTTCCACCTGGTGCAGATGATGGGTGGAATGCCAGCTC--A	5449
Db	1142	snIleProProHisIleGlyValProMetGlyProGlyIleuMetSerHisAsnProIleuMetG	1162
QY	5450	ATTTTGGTTCCGAATTTCAATCCGAATATT--CAGGTAAAGCGAGTACCCCAAC---	5502
Db	1162	lyHisGlySerGlnIleuProProMetValProGlnGlyArgMetGlyIleuProGlnIleu	1182
QY	5503	-----ACATCAATACATACATCCAGATGAGGAGCAGACAGCCAAACATTAACATG	5557
Db	1182	heProProValGlnSerProProGlnGlnValAlaProIleProHisAsnIleYProSerGlyG	1202
QY	5558	GAGCTAATTAATGTGCGAAATGCCACTAGCTGGAAATTTTTCAGAGGTACGCTAACCTC	5617
Db	1202	lyGlnGlySer-----PheProGlyGlyMetGlyPhe-----ProG	1214
QY	5618	AAATGGGAGCTGTAGACATGGGTGGCCAAATATGCCCAATACAGACGACGCTACTC	5677
Db	1214	lyGlnGlyProIleuIleYArgProSerAsnIleu--ProGlnSerSerIleAspAlaAlaL	1233
QY	5678	CTGGAATGCCAGATTTGATGGCGGAGCAGGACCGGAGGATATGCTAATTTCT	5733
Db	1233	euCylArgPProGlyIleYProGlyGlyProAspSerPheThrValIleuGlyIleuSerMetP	1253
QY	5733	-----	5733
Db	1253	roSerValPheThrAspProAspIleuGlnGlnValIleArgProGlyValaThrGlyIleP	1273

Oy		5733	-----	-----	-----	5733							
Dd		1273	roglupeapleusefargllelprosergdluyprosergdlntheunglntyr			1293							
Oy		5734	----TCCGAGACCAACCACAGAACAATGCACAAACAATTCTGGGGCAAGCATGTGA			5788							
Dd		1293	heProArGdlYglUvalIProGlYArgylseGIInPrOgInglyPro-----GlyP			1311							
Oy		5789	TTAACTCCTTTCCGAATTTGCCAATTCMAATGCTACTATTGTTGACGAAGAGGR-			5838							
Dd		1311	heserHismetGln-----GlymetMetGlyglUgnlalProArghMetC			1326							
Oy		5839	----GGATTAACCGGCATGACCGATCAATGATAATGTGTCAACCATCTATGATTAAGG			5893							
Dd		1326	lyLeualalaleuProGlYmetGlyglYProglYProvalGlyltmrProAsplleroleug			1346							
Oy		5894	GCAATGCGTCCA-----CATGCCATGCGGCCAAT-----GTAATGGSGTG			5932							
Dd		1346	IyThralaProSerMetProGlYhlIsanProMetaGrProprolaPheleunglng			1366							
Oy		5933	CQGCGATGCCAACCGGTAAACAGGAATTCAGTTTGACAGTCATCGATGRTATGACT			5992							
Dd		1366	lyMetMeGlgYProhIHierghMetseterProAlagInsersetThr-----M			1382							
Oy		5993	GTGTGCGGGATCCGTCATCAATTTTTCTACTAAGCTTCTGCCAACAGCGCTGGACACACA			6052							
Dd		1382	eCProgIyInPrOthrleumetSerAnProAlalaIAla-----			1395							
Oy		6053	TGTTTGATGACGACCAACAGGCCAATGACGCTTAAGACACAACATTAACCTTA			6112							
Dd		1396	-----ValGlymettlePro-			1400							
Oy		6113	GTGGAAATGTGTCAAACAACCAATCCGGACTTCAGTAGGACAAAGGCGACATCCACTGCATG			6172							
Dd		1401	--GlylvsaAParGlyProIacetyLeuYrTrhhIsPRoGlyProvalGlySerProg			1420							
Oy		6173	GGCAAGACATGCCGACGGGTCAAGCTTTAATGTGAACTTACTAATTAAT			6222							
Dd		1420	lyMetMetMetSerMet--GlnIlymetMetGlyProAdenArgrThSer			1435							
<hr/>													
RESULT_8													
ID	ADQ18945	standard; protein;	1394 AA.										
XX	ADQ18945;												
AC													
XX													
DT	26-AUG-2004	(first entry)											
XX													
DE	Human soft tissue sarcoma-upregulated protein - SEQ ID 1764.												
XX													
KW	soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.												
XX													
OS	Homo sapiens.												
PN	M02004049838-A2.												
XX													
PD	10-JUN-2004.												
XX													
PF	26-NOV-2003; 2003WO-USO38193.												
XX													
PR	26-NOV-2002; 2002US-0429739P.												
XX													
PA	(PROT-) PROTEIN DESIGN LABS INC.												
XX													
PI	Aziz N, Gineburg WM, Zlotnik A;												
DR	WPI; 2004-441208/41.												
PT	Early detection of soft tissue sarcoma comprises determining expression												
PT	of a gene in a first soft tissue sample and a normal soft tissue sample												
PT	and comparing the gene expression, also useful in treating soft tissue												
XX	sarcoma.												
XX													

PS Example 2; SEQ ID NO 1764; 210bp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX Sequence 1394 AA:

Alignment Scores:

Pred. No.:	3,12e-17	Length:	1394
Score:	314.00	Matches:	306
Percent Similarity:	30.12%	Conservative:	172
Best Local Similarity:	19.28%	Mismatches:	509
Query Match:	2.61%	Indels:	601
DB:	8	Gaps:	70

US-10-664-859-1 (1-6909) x ADQ18945 (1-1394)

QY 2397 AAAATTGAAAGA---ATTTCAGACAGATACACGAAAAAAGATGCTCTTGACA 2453
DB 98 LysArgGlnArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr----- 115
QY 2454 ATGAATATAGCAAGAAAGACATGAGAGCTGCAATGATGATCCGATTTATCAAT 2513
DB 116 ---ProAsnAspAspSerAspIleLeuGlyGluSerAsnSerAlaAspHisIleLeuSerGln 134
QY 2514 GAATCTTAATATATCTCTGCAATTCGAGCATATTAGAGCGAGAGACCAATACC 2573
DB 135 AspSerGlnHisThrProHisSerMetThr----- 144
QY 2574 GGAATCGAGAGTGGAGCGGAGCGGAAATTTATGACGCCAAGCCAAATGGAATCC 2633
DB 145 -----ProSerAsnAlaThrAlaProArgSerSerThrProSer 157
QY 2634 TCGGATGACGATATTGTTGATTACATGACAGCAAAATCAC-----ATA 2681
DB 158 HisGlyGlnThrThrAlaThrGlnProThrProAlaGlnThrProAlaValVal 177
QY 2682 TTGGTGTTCACATCTGAGCGCAACAAAGGGCGCAATCAGTTTAAGGGGTCAATT 2741
DB 178 TyrValPheSerThrGlnMetAlaAsnLeuAlaHisGlnValValLeuValGlyGlnVal 197
QY 2742 CAAACTATTAATGCGTATCACTGCACTCAGCTGCTCAAAAAGCTTCTGGAAGACTTT 2801
DB 198 GluThrIleValSerPheHis----- 204
QY 2802 TTTATGAAAAACCTTTAAAGATTACAGTTACAGCGGCACAAATTCGTCGATGCCA 2861
DB 205 ---IleGlnAsn---IleSerAsnAsnLeuThrGlnArgSerThrAla-----Pro 219
QY 2862 TGGATAGCATGGGGGAGTTGACATACTCTCTATCTCTGAGCCAAATAACAA 2921
DB 220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLeuProLeu-----Pro 235
QY 2922 CAGCAGCCACATCAAAAGACCGTAGGCTATTGAAACCCCAATTCACATCAATAAAC 2981
DB 236 GlnGlnProProValProAla-----AsnGlnAspGlnAsn 247
QY 2982 AGCAAGTACTGATGAGCGCGCTAGCAACTCTTT----- 3020
DB 248 SerSerGlnAsnThrArgLeuGlnProThrProThrProAlaProAlaProLeuPro 267
QY 3021 -----GTCGACCACTGATCTGATCTATGGCAACGAAATGAAATTTGATG 3062

DB 268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLeuLeuIle 286
QY 3063 TGGTGGAAAGCGGATCTTCAACACACAGI-----AGTCTGCAACAAAC 3107
DB 287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306
QY 3108 TCACGAATATCATATAGACATATCAGTATCCAGCAGGACGACAGCAATTAAGATCTG 3167
DB 307 SerThrProAsnAsnArgAlaValThrProValSerGlnIleSerAsnSerSerAla 326
QY 3168 GAACAGCTGGCGCTTGATTTGGACAGCTCAAAAAGAACGATCCTGAGCTGACACT 3227
DB 327 AspProValAlaProProProProProValSerSerGlyGluProThrLeu---Gly 345
QY 3228 GAAAAACAATTGTATCTACGACAGAGTTAAGTTCAAGCAACAACTTACACCA 3287
DB 346 GluAsn-----ProAsp---GlyLeuSerGlnGlu 354
QY 3288 CAGCGGCAATCGGGAGAAACAGTTGGCAAAAATAAATAAATCAATTTCTTTT 3347
DB 355 GlnLeuGlnHisArgGlnArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374
QY 3348 CCGTGAATAAGAAATTCAGTAGAGCTTAATGTAAGCTCACAGATCAAAAATTCACAGA 3407
DB 375 ProAspGlnLeuGlnPheThrGlyAlaGln----- 384
QY 3408 GATTTAATGATGGGATGTGGGTGGCGAGGGGAGATCTTTAATTCGACAGATGCA 3467
DB 385 -----SerGlyGlyProGlnGlnAsnPro----- 392
QY 3468 CAATGATATGCGAGGTAACGCCAATCGAGCTTATCGCGCAAGTTGACGACTT 3527
DB 392 ----- 392
QY 3528 TCGGAAGATGTAATGCATCCAGGAGATGTAATATCAGATATGGTCCGTAATAGATG 3587
DB 393 -----GlyValLeuAspGlyPro 398
QY 3588 AATTAATATCAAAAACCAAGTGCATGTCGATCTGGAATAGTGTTCACCTGGAACA 3647
DB 399 GlnLeuLeuSerProGlnGlyProIleGlnAla----- 408
QY 3648 ACTGCAGCTGAGTAAATGATCATATGCAATGCTCAAGTCCGCGCCCGCAATGGCAAT 3707
DB 409 -----MetMetAlaGlnSerIleSerLeuGlyValGly--- 419
QY 3708 ATGATGGAAAGCTCTACGATATGCTAGCTCTGTTGGCAACACAACTGCAACGTATC 3767
DB 420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlnHisArgAspVal 438
QY 3768 GGAAGGCCCCAGAT--ATGCTTAAGAAATTTAATCAAGATACCCGAAAC----- 3818
DB 439 ProPheSerProAspGlnMetValProProSerMetValSerGlnSerGlyThrIleGly 458
QY 3819 ---CATTCACATCAAGGGGAGTTGCTCAATAGAGGTGCAAGATTCACATCA 3872
DB 459 ProAspHisLeuAspHisMetThrProGlnGlnIleAlaThrLeuValSerGlnGlnGlu 478
QY 3873 TTTTTCGAAGAGCGCTCAAGGGGGCGCAAGCCGACAGCACTGCACTGAGTATGACCA 3932
DB 479 PheTyrGlnGlnLeuValArgGly-----LysProGlnGlnVal-----ValValGln 493
QY 3933 CAA-----CAGCAAAACCTTCTGATCTGATCTGATCTGATCTGATCTGATCTG 3965
DB 494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly----- 509
QY 3966 TCGTTAAACAACAGGTGCAAGCCCTGCAAGTCCACTCTCTTACAC----- 4016
DB 510 -----ValValArgGlyProProProProProProProProProProProPro 522
QY 4017 -----TTCATCCAGAGATCTGCG 4034

Db 523 SerGluGlyTTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
QY 4035 TCAGTACCAATAGCCACTGATCCCATCCCTCGAGTCCAAACATCTA----- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
QY 4086 -----TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
QY 4104 ACAACCGACAGCATGATGGATTG-----CCGACCAACTCTCTCGACATG----- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTTrpProAspPheValProAlaSerProAspGly 602
QY 4149 -----GATGGACAGCATCATCTCTGA----- 4172
Db 603 ArgAsnPheProProGlyArgGlyIlePheSerGlyProGlyArgGlyIleArgPhePro 622
QY 4173 -----TCTGTTCGGCAAGCTTAATCTTCGACGGTTCAGGACGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluGlySerGlnLeuGly 642
QY 4212 ACAACA-----ACAGTCTCTCAGCAACAAGACTGTTTCAGGACAGACCCCATG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
QY 4266 CCGTCAATCAAAATCGTAGTAAGAAATACCGATCCGATCCGTTCTTACGCATTACTTA 4325
Db 663 ProGlySerGln--ArgHisMetGluProGlyAsnAsnProIlePheProArgIlePro 681
QY 4326 AGCAGCAACCAAGTACCCCTTATCTCATCTATCCCAAG----- 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProGlyIleProGlnMet 701
QY 4368 -----GAATTGAGTCTTTCGTCAGTCCCTCGCTGCTGATTAATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu--PheGlyMetValProSerGlyMet----- 716
QY 4416 AATTTTTTAAAGACAATCAAAATATGAAATATGCTTAATTAATTAATTAATTAATTA 4475
Db 717 -----LysGlyAspValAsnLeuAsn--ValAsn----- 725
QY 4476 CTGCGAAATTTGATAGAAAAATCAGGAATAGAAAAATTAATTAATTTCCGACCG-- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnIleMetArgIleAlaGlyIleGlyProGlu 745
QY 4532 ----- 4532
Db 746 GluMetLeuLeuLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnIleMetVal 765
QY 4533 -----CCCATCCATTTCTTG 4547
Db 766 ProLeuProPheGlyGlnHisProGlnGlnGluTyArgIleMetGlyProArgProPheLeu 785
QY 4548 -----AATCAATTTCTGAGTATTTGATAGATAA----- 4579
Db 786 PrometSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
QY 4580 -----TCTACTATTAATAATTAACAGAAATTCATATCCGTTAATTTGAAGAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu----- 822
QY 4628 CACTATTGTTTAATTAAGAAATTAATAATATATTATTAATATTTCTACAGGTGATAC 4687
Db 823 -----AsnProSerSerAsn 827
QY 4688 ATGAAAGTGGGACCAAGCCACAGGCTCAGCGGTACACAGTAATTAAGTCTTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg-----GlyLeuGlyArg 843
QY 4748 GCAAAATTAAGATGATGATTTGCTGATCC-----AGTCTGGTTTAAACCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro-- 862

QY 4799 CCACATATGCAAGAACATTCAAATTCAGCATTAACGCTATTAATAGGCTCTACCAAT 4858
Db 863 -----LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
QY 4859 ATACGATGAGAGTAAATATTTAAATATTTTATTAACGTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu----- 883
QY 4919 TTTTTCAGCGTCAGATCAGCGCAAGGTGATCCGTATCAATTTAGTCCGC----- 4969
Db 884 -----LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
QY 4970 -----GCTCCGATAATATTCGCTTAATCCCAATAGCGCAATCGCGCC----- 5015
Db 898 GlyProAlaAlaAlaLaserIleLysSerProProValLeuGly--SerAlaAlaLase 917
QY 5016 -----ACCAACACAA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTTrpThrSerSerP 937
QY 5027 TGACCAAAACTTCGATCCATCTCTTTGGACAAATGTCACCAACTPAACAAGTT 5086
Db 937 oLysProProLeuGlnSer-----ProGlyIleProProAsnHisLysAl 952
QY 5087 GCGTCCAGCATGGGTAGTCCAGCCGAACGTGGTGTATGACAGATGAGGGAGTCCG- 5145
Db 952 aProLeu--ThrMetLaserProAlaMetLeuGlyAsnValGlnSerGlyGlyProProP 972
QY 5146 -----GGACCGTCGACATCAATTTGAG----- 5169
Db 972 roProThrLaserGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrp 992
QY 5169 ----- 5169
Db 992 roTyThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
QY 5170 -----CATGGAATTAATTTCCGGAC 5188
Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuThrHisAspAlaIleLysThrv 1032
QY 5189 TAGATGATCAGGAATAGTATACC-----ATAATCAAAATAACTGTCAATTAAG 5242
Db 1032 aAlaLaserSerAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051
QY 5243 TCGTAATGAATCAATGGGTCCCGCAATGCTGAATCTTAATAATGTGCTAGCAGCGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle-----SerGlyP 1068
QY 5303 CAAT----- 5307
Db 1068 roAsnProValAlaProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
QY 5308 -----GACCGCGCTGCTTAATCTTAATCTCAATCCCAATGGTGATTA 5350
Db 1088 iAsnArgIleMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
QY 5351 GAGAGATTCATAGGCTCGTGGTGGCTCAGCAACTTCAACTTTCAAGGGGTG 5410
Db 1108 -----PrometGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
QY 5411 TTCACCTGGTGCAGAAATGATGGGTGGAATGCCAGTCAATTTGGTTGAAATTTCAATC 5470
Db 1125 lngluProProMetValProGlnGlyArgMet-----GlyPheProGlnGlyPheProP 1143
QY 5471 CGAATTTTCAAGTAAAGGAGTACCCCAACCACTACAATATACATCCAGTAAGGCAC 5530
Db 1143 ro-----ValGlnSerProProGlnGlnValP 1152
QY 5531 AGAAGCCACAACAATAACAAATGAGCTAATTAATGTGGAATGCCACTAGTCTGG 5590
Db 1152 roPheProHisGlnGlyProSerGlyGlnGlySer-----PheProGlyGlyMetG 1170

OY	5591	AATTTTTCACAGAGGTAACGCTTAACCTTCAAAATGGGTCGCTGTAAGGCAATGGGTGCGCAATAT	5650
Db	1170	lypne-----ProGlyGlnGlyProLeuGlyArgProSerAsnLeu-	1183
OY	5651	GCCCAACATCATGACCGACGACGGTACTCTCGGAATGCGAGATGATGGCGGAGACGAGG	5710
		::::	
Db	1184	--ProGlnSerSerAlaSerAlaAlaLeuCylybysProGlyGlyProGlyGlyProAsp	1203
OY	5711	CCGGAGGTATGCTTAATGAATCT-----	5733
		:::	
Db	1203	erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnIuv	1233
OY	5733	-----	5733
Db	1223	allLeArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG	1243
OY	5734	-----TCGGAGAGCAACACAGAAACAGATCA	5761
Db	1243	IuLybProSerGlnThrLeuGlnIlyrPheProArgGlyGlyIuValProGlyArgIyGlnP	1263
OY	5762	CAAAACAATCTCGGGGCAAGCAATGGTATTAATCTTTACAAATTCAGAAATGTCTA	5821
Db	1263	roGlnGlyProGlyPro-----GlyPheSerHisMetGln-----GlyM	1276
OY	5822	TTGTTCAGCAGAGAGGGT-----GGATTACCCGGCGCATACGGATCAATGA	5866
		::::: ::::	
Db	1276	etMetGlyGlnGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP	1296
OY	5867	ATATTGGTCACCATCTATGTGATTAAGGGGCGATGCGTCCA-----CATGCCA	5911
		:::::	
Db	1286	roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM	1316
OY	5912	TGCGGCCCAAT-----GTAAATGGGTGCGCGAGATGCCACCGTTAAACGCGAAATTCAGT	5965
		::::	
Db	1316	eArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisIleArgMetMetSerP	1336
OY	5966	TTGCACAGTCATTCGGATGGTATTTGACTGTGTGGGGATCCGTCATCATTTTTCACCTAACG	6025
		::::	
Db	1336	roAlaGlnSerThr-----MetProGlyGlnProThrLeuMetSerAsnProA	1352
OY	6026	CTTCTCGCAACAGCGCTGGACACACATGTTTGATACGACACAAACGCGCATACGCTA	6085
		::::	
Db	1352	IaAlaAla-----	1354
OY	6086	AGACACACACATTAAGAACAATACCTAGTGGAAATGTGCAAAACCAATCGGGAATTGCAG	6145
		::: ::::	
Db	1355	-----ValGlyMetIlePro--GlyLybAspArgGlyProAlaGlyLeuIlyrT	1370
OY	6146	TGGACAAAGGCGACATCCAACTGCATGGGCAAGACATGCGCAGGGTCAGTCTTAATG	6205
		::::	
Db	1370	hrHisProGlyProValGlySerProGlyMetMetMetSerMet--GlnGlyMetMetG	1389
OY	6206	GACCTACTAATTAATAT 6222	
		::::	
Db	1389	IyProAsnArgThrSer 1394	
		::::	
RESULT 9			
AAB71229			
ID	AAB71229	standard; protein; 1426 AA.	
XX	AAB71229;		
DT	18-NOV-2002	(first entry)	
DE		Human legless homologue lgs/bcl9 protein.	
XX			
KW		legless; human; lgs; wnt/wingless signaling pathway; wnt; wg;	
KW		tissue proliferation; tumour; cystostatic; cellular disorder; colon;	
KW		blood disorder; cancer; breast; head and neck cancer; brain; thyroid;	
KW		medulloblastoma; skin cancer; tissue regeneration; tissue repair.	
OS		Homo sapiens.	
XX			

[illegible]

Db 766 ProleuProhneGlyLinhierProGlnGlnIuYrGlyMetGlyProIatxProhneleu 785
 Qy 4548 -----AATCCAAATTTGAGATGATTTGTAAGATAA----- 4579
 Db 786 PrometSerGlnIyProGlySerAmsnserGlyLeuAargAanLeuAargIuProIleGly 805
 Qy 4580 -----TCCACTATTAAATTAACACGAAATTCATATCCGTTATTTGAAAT 4627
 Db 806 ProAerGlnAargIuAmsnserAargLeuSerHiSmetProProleuProleu----- 822
 Qy 4628 CACTATTGTTTAATTAAGAAATTAATAATATGTTTATTAATATTTCTACAGTGAATAC 4687
 Db 823 ----- 4687
 Qy 4688 ATGAAGTAGGCGCACCAAGCCACAGGCTACAGGCTACCAAGTAATATGTTATATAG 4747
 Db 828 ProTherSerLeuAnthrAlaProProValGlnArg-----GlyLeuGlyArg 843
 Qy 4748 GCAAAATTAAGATGATGATTTGTCATCC-----AGTCCGTTTAAACCCGAT 4798
 Db 844 LybProleuAerPiIleSerValAlaGlySerGlnValHisSerProGlyIleAmsPro-- 862
 Qy 4799 CCACATATGCAACCAATTCAAATTCAGATTAAACGCTATTAATGAGCTTACCAAT 4858
 Db 863 -----LeuYsSerProTherMechiSglnValGlnSerProMetLeuGlySerProSer 880
 Qy 4859 ATACAGATGAGGTAATATTTAAATTTTAAATTTTAAACGTTTGTGTTAATTTATCTTC 4918
 Db 881 GlyAmsLeu----- 883
 Qy 4919 TTTTTCAGCGTCAGCATCAGCGAGGTGATCCGTACATTTAGTCGC----- 4969
 Db 884 -----LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
 Qy 4970 -----GCTCCGATAATATTCGCTAAATCCCAATAGGCAATGCGCGCC----- 5015
 Db 898 GlyProAlaAlaAlaAlaSerIleYsSerProProValLeuGly-SerAlaAlaAlaAla 917
 Qy 5016 -----ACCAACACAGA 5026
 Db 917 rProValHiLeuYsSerProSerLeuProAlaProSerProGlyTrpThrSerSerP 937
 Qy 5027 TGACCCAAACTTGATGCCATCTCTTTGGACAAATGCCCAACACTTAACAAGTT 5086
 Db 937 OlybProProleuGlnSer-----ProGlyIleProProAmsHiSlyVal 952
 Qy 5087 GCGGTCCAGCATGGTATGATCCAGCCGAACTGGTGATAGCATGATGAGGAGTCCG 5145
 Db 952 aProleu-ThrMetAlaSerProAlaMetLeuGlyAmsValGlnSerGlyIyProProP 972
 Qy 5146 -----GACCGTCGACATCAATATTGAG----- 5169
 Db 972 roProThrAlaSerGlnProAlaSerValAmsIleProGlySerLeuProSerSerThrP 992
 Qy 5169 ----- 5169
 Db 992 roTyThrMetProProGlnProThrLeuSerGlnAmsProLeuSerIleMetMetSera 1012
 Qy 5170 -----CATGAATATTTCCGGAC 5188
 Db 1012 rGmetSerLysPheAlaMetProSerSerThrProleuTyThrAlaPheAlaIleLysThrV 1032
 Qy 5189 TAGATGATCAGATAGATAC-----ATMAATCAAAATACTGTTCATTCATGATG 5242
 Db 1032 AlAlaSerSerApsApsApsSerProProAlaArgSerProAmsLeuProSerMetAms- 1051
 Qy 5243 TCGTAAGACTCAATGGTCCCGAATGCTGATCTGAATCTTAAATGCTGCTACAGCGCTC 5302
 Db 1052 --AmsMetProGlyMetGlyIleAmsThrGlnAmsProArgIle-----SerGlyP 1068
 Qy 5303 CAAAT----- 5307

Db 1068 roAmsProValAlaProMetProThrLeuSerProMetGlyMetThrGlnProleuSerH 1088
 Qy 5308 -----GACCCGCTGAGCTTAATCTTAATCTCCCAATGCTGATTA 5350
 Db 1088 lssAerGlnMetProSerProAmsAlaValGlyProAmsIleProProHiSgIyVal- 1107
 Qy 5351 GAGAGAAATTCATGAGGTGCTGCTGCTGCTGAGCAATCTTCAAACTTTCAAGGGGTG 5410
 Db 1108 -----PrometGlyProGlyLeuMetSerHiSamsProIleMetGlyHiSgIySerg 1125
 Qy 5411 TTCACCTGTGTCAGATGATGGGTGAATGCGCATGATTTGTTGTTGATTCATC 5470
 Db 1125 lngIuProProMetValProGlnGlyArgMet-----GlyPheProGlnGlyPheProP 1143
 Qy 5471 CGAATATTCAGGTAAAGCCAGTACCCCAACACATATACATATCCAGTAAGGGCAC 5530
 Db 1143 ro-----ValGlnSerProProGlnIlnValP 1152
 Qy 5531 AGAAGCCAAACAAATTAACAAATGAGCTAATATGTCGAATGCGACCTAGTCTG 5590
 Db 1152 roPheProHiSamsGlyProSerGlyGlnGlySer-----PheProGlyIyMetG 1170
 Qy 5591 AATTTTGACAGAGTACGTAACCTCAAAATGGGTGCTGATGCAATGGGTGCGCAATAT 5650
 Db 1170 LyPhe-----ProGlyGlnGlyProleuGlyArgProSerAmsLeu- 1183
 Qy 5651 GCCCAACATCAGCCAGGAGCGTACTCTGGAATGCCAGATTTATGCGGAGACAGAG 5710
 Db 1184 --ProGlnSerSerAlaApsAlaAlaLeuCySylsProGlyGlyProGlyIyProAmsP 1203
 Qy 5711 CCGAGGATGCTAATGATTTCT----- 5733
 Db 1203 erPheThrValLeuGlyAmsSerMetProSerValPheThrApsProApsLeuGlnIuV 1223
 Qy 5733 ----- 5733
 Db 1223 allIeArgProGlyAlaThrGlyIleProGlnPheApsLeuSerArlIeIleProserG 1243
 Qy 5734 -----TCCGAGAGCAACACCAAGAACAGATCA 5761
 Db 1243 lulybProSerGlnThrLeuGlnTyrrPheProAargIyGluValProGlyArgIySglnP 1263
 Qy 5762 CAACATCCTGGGCGCAAGCAATGCTATTAATCTTTCGAATTCATCAATATGCTA 5821
 Db 1263 roGlnGlyProGlyPro-----GlyPheSerHiSmetGln-----GlyW 1276
 Qy 5822 TTGTTGACGAAGAGGT-----GGATTACCCGCGCATGCGATCAATGA 5866
 Db 1276 etMetGlyGlnGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyIyProGlyP 1296
 Qy 5867 ATATTGTTCAACCTATATGATAAGGCGCATGCTCA-----CATGCCA 5911
 Db 1296 roValGlyThrProApsIleProleuGlyThrAlaProSerMetProGlyHiSamsProm 1316
 Qy 5912 TCGGCGCAAAAT-----GTATGGGTGCGGCGGATGCCACCGCTTAACAGGCAATTCAGT 5965
 Db 1316 etArgProProAlaPheLeuGlnGlnIyMetMetGlyProHiSamsArgMetSerP 1336
 Qy 5966 TTGACACAGTACGATGATGATTAAGATGCTGCGGAGTCCGTCATCTTTTCTCAACG 6025
 Db 1336 roAlaGlnSerThr-----MetProGlyGlnProThrLeuMetSerAmsProA 1352
 Qy 6026 CTTCCTGCAACAGCGCTGACACACATGTTTGATGACACAAAGCCCAATCAGCCTA 6085
 Db 1352 lAlaAla----- 1354
 Qy 6086 AGACACAAACATPAAGAACATACCTAGTGAATGTGTCAAAACCAATCGGACTTGCAG 6145
 Db 1355 -----ValGlyMetIlePro--GlyYsApsArlGlyProAlaGlyLeuTyrr 1370
 Qy 6146 TGGCAAGAGGAGCATCACTGATGGGACAGAGACATGCGACAGGTCACTTTAATG 6205
 Db 1370 hrHisProGlyProValGlySerProGlyIyMetMetSerMet--GlnGlyMetMetG 1389

QY	6206	GACCCACTCAATTAATTA-----	-----	ANGTCAACGCCGGAAGTCAAGTCA	6253
Db	1389	lypProGlnGlnAsnIleMetIleProPogIImetAArgIy---	-----	MetAlaIaA	1408
QY	6254	CTAACGATGCTCTCGCATCAATTCGAGGTCCTCTCT	6294		
Db	1408	spValGlyMetClyIyPheSer---GlnGlyProGlyAsn	1420		
RESULT 10					
ABW01534					
ID	ABW01534	standard; protein; 1426 AA.			
AC	ABW01534;				
XX					
DT	15-JAN-2004	(first entry)			
XX					
DE	Human IgG/bc19 protein.				
KW	Legless protein; IgG; cell fate disorder; blood disease; gene therapy;				
XX	cancer; tissue regeneration; tissue repair; cytostatic.				
OS	Homo sapiens.				
XX					
PN	US2003114413-A1.				
PD	19-JUN-2003.				
XX					
PF	19-DEC-2002; 2002US-00322579.				
XX					
PR	28-JUL-2000; 2000US-0221502P.				
XX					
PR	27-JUL-2001; 2001US-00915543.				
XX					
PA	(UYZU-) UNIV ZURICH.				
XX					
PI	Basler K, Brunner E, Froesch B, Kramps T, Peter O;				
XX					
DR	WPI; 2003-829432/77.				
XX					
DR	N-PSDB; AAD62642.				
XX					
PT	Novel IgG polypeptide useful for isolation of IgG-binding proteins,				
XX	diagnosing disorders of cell fate, treating diseases such as cancer.				
PS					
XX	Example 2; Fig 8B; Opg; English.				
CC					
CC	The invention relates to novel legless (lgs) proteins and polynucleotides				
CC	encoding such proteins. Lgs sequences are useful for the treatment of				
CC	disorders of cell fate such as differentiation or proliferation. The				
CC	invention is used to treat blood disease or a cancerous condition				
CC	characterised by over-stimulation of the Wnt pathway such as colon,				
CC	breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and				
CC	is administered to prevent progression from a pre-neoplastic or non-				
CC	malignant condition to a neoplastic or malignant state. It is				
CC	administered to promote tissue regeneration and repair. The invention is				
CC	also useful in the therapy of diseases cost by an over-activation of Wg				
CC	pathway. It is useful for reducing lgs gene expression in an invertebrate				
CC	or vertebrate organism or an invertebrate or vertebrate cell line. The				
CC	invention is also useful in gene therapy. The present sequence is human				
CC	IgG/bc19 protein used in the invention				
XX					
XX	Sequence 1426 AA;				
Alignment Scores:					
Pred. No.:	4.69e-17	Length:	1426		
Score:	312.00	Matches:	313		
Percent Similarity:	30.34%	Conservative:	177		
Best Local Similarity:	19.38%	Mismatches:	519		
Query Match:	2.59%	Indels:	607		
DB:	7	Gaps:	73		
US-10-664-859-1 (1-6909) x ABW01534 (1-1426)					
QY	2397	AAATTTGAAAGC---ATTTCAAACGACAGTACACGAGAAAAAAGATCGCTCTTGACA	24533		

[illegible]


```

385 -----SerGlyProGlnIleAsnPro----- 392
3468 CAACTGATATGCGAGTAAAGCAATCGAGCTCTTATCGCGACAATTGACAGACT 3527
392 ----- 392
3528 TCGGAAGATGTAAATGATCCAGGGAGTGTATATCAGATATGGTCCGTAATAGAGATG 3587
393 -----GlyValIleAsnArgGlyPro 398
3588 AATTAATATCAAAAAACCAATGTCATGTCGATCTGGAGTAGGTGTGTCACTGGACAA 3647
399 GlnIleValSerProGlnIleProIleGlnIleA----- 408
3648 ACTGCAAGCTGAGATTAATGTCATATGTCATGTCGCAAGCTCCGGCCCCCGAATGGCAAT 3707
409 -----MetMetAlaIleSerGlnSerIleGlyValGly--- 419
3708 ATGATGGAAAGCTCTACGATATGTCATGTCGTCCTGTTGGCAACACAAAGCTCAAGTCATC 3767
420 ---ProGlyProArgThrIleAsnValGlyAlaProPheGlyProGlnIleGlnIleValGlyAlaProVal 438
3768 GGAACGCCCCAGAT---ATGCTTAAGAAAGTTTAAATCAAGATACCCGCAACC----- 3818
439 ProPheSerProAsnGlnMetValProProSerMetAsnSerGlnSerIleGly 458
3819 -----CATTCATCATCAAGGGGAGTGTCTCAATAGAGTGTGCGAAGATTCAAACATCA 3872
459 ProAsnIleValAsnIleAsnMetThrProGlnIleIleAlaThrIleValSerGlnIleGlnIle 478
3873 TTTTTCGAAGAACGCTTCAAGGGGGGAGCCGACGACCAAGTCACTGGAACTGTAGTACCA 3932
479 PheTyrlGlnIleValAsnArg---LysGlnIleGlnIleVal-----ValValGln 493
3933 CAA-----CAGCAAAACCCCTTCTGAGTGTGTGAAC 3965
494 GlnCysSerIleGlnIleAsnMetCysValIleGlnIleGlyProArgIle----- 509
3966 TCGTTAAACACACAGGTCGACCCCTGCAAGGTCGACCTCTCTTATCCAC----- 4016
510 -----ValIleValArgIleProProProProGlnMetThrPro 522
4017 -----TCCATCCAGAGATCTGCG 4034
523 SerGlnIleYTrpAlaProGlnIleGlyThrGlnProPheSerAsnArgIleIleAsnMetProIle 542
4035 TCAGTACCAATAGCACTCAATCGCCCAATCCTCGAGTCCAAACATCTA----- 4085
543 SerIleuProProArgIleMetAlaProIleProIleAsnMetProGlySerGlnMetArgIleu 562
4086 -----TCTCTCCCGTCAACCGCGG 4103
563 ProGlyPheAlaGlyMetIleAsnSerGlnMetGlnIleProAsnValProAsnProAla 582
4104 ACAACGCAAGCATGGATG-----CCGACCAACTCTCTGACATG----- 4148
583 SerArgProGlyIleuSerGlyValSerTyrProAsnProValProLysIleProAsnArgIle 602
4149 -----GATGAAACAGATCATTAATCTGA----- 4172
603 ArgAsnPheProProGlnIleGlyIlePheSerGlyProGlyArgGlyIleArgPhePro 622
4173 -----TCTGTTCCGCAAGCTAATCTTCGACGTTCCAGGCGC 4211
623 AsnProGlnIleuSerGlnIleuMetPheGlnIleGlnIleuAlaGlyLysGlnIleGly 642
4212 ACAACA-----ACAGTGCCTCAGCAACAACAAGATGTTTTCAGGACGACACCCCATCG 4265
643 LeuProProGlyIleMetAlaMetGlnIleArgProSerMetGlnMetAlaAsnMetIle 662
4266 CGCTCAATCAAAATGTAAGTAATAACCGATCTGCAACGTTCTTACGATTAATCTTA 4325
663 ProGlySerGln---ArgHisMetGlnProGlyAsnAsnProIlePheProArgIlePro 681
4326 AGCAGACACCAAGTACCCCTTATCTCATCTATCCCAAG----- 4367
682 ValGlnIleProIleuSerProSerArgGlyAsnPheProLysGlyIleProProGlnMet 701
4368 -----GAATTGAGTCTTTCGGTCACTCTCTGCTGGATGTTATATTGTTT 4415
702 GlyProGlyArgGlyIleuGln---PheGlyMetValProSerGlyMet----- 716
4416 AATTTTAAAGACAAATCAAAATATGAATGTGGTTAATTAATTAATTAATTAATTA 4475
717 -----LysGlyAsnValAsnIleAsn---ValAsn----- 725
4476 CTCGAAATTTGATCAAAAAAATACAGAAATAGAAAAATTAATTTATTTCCGACCG--- 4532
726 MetGlySerAsnSerGlnMetIleProGlnIleMetAlaGlyIleAlaGlyIleProGln 745
4532 ----- 4532
746 GlnMetLeuIleValArgProGlyIleSerAsnMetLeuProAlaGlnIleGlnIleMetVal 765
4533 -----CCCATCCATTCTTG 4547
766 ProIleuProPheGlyIleIleAsnProGlnIleGlnIleValArgIleMetGlyProArgProPheIleu 785
4548 -----AATCCAAATTTCTGAGTGTATGTAAGATAA----- 4579
786 ProMetSerGlnIleProGlySerAsnSerGlyIleuArgAsnIleuArgGlnProIleGly 805
4580 -----TCTACTATTAATAATTAACACGAAATTCATATCCGTTAATTAATAAT 4627
806 ProAsnGlnArgHisAsnSerArgIleuSerHisMetProIleuProIleu----- 822
4628 CACTATTGTTTAATAAGAAATTAATAATATGTTTATTAATATTTCTACAGGTGATAC 4687
823 -----AsnProSerSerAsn 827
4688 ATGAAGTAGAGGACACAGCCGACAGGGTCAGCGGTCAACCAATTAATATCTAATAGAG 4747
828 ProThrSerIleuAsnThrAlaProProValGlnArg-----GlyLeuGlyArg 843
4748 GCAAAATTAAGATGTACGATTTGCTGCATCC-----AGTCTGGTTTAAACCGCAT 4798
844 LysProIleuAsnIleSerValAlaGlySerGlnValIleSerProGlyIleAsnPro--- 862
4799 CCACATATGCAAGCAATTCAAATTCAGATTAAGCCCTATTAATAATGGCTTACCAAT 4858
863 -----LeuIleuSerProThrMetCnIleGlnValGlnSerProMetIleuGlySerProSer 880
4859 ATACAGATGAGGTAAATATTTAATTAATTTTAAAGTTTGTGTTAATTTATCTTC 4918
881 GlyAsnIleu----- 883
4919 TTTTTCAGCGTCAAGCATCAGCCCAAGTGCATCCGTACATTTAGTCGC----- 4969
884 -----LysSerProGlnIleThrProSerGlnIleuAlaGlyMetLeuAla 897
4970 -----GCTCCGATTAATATTCGCTTAATCCCAATGTCGCAATGCGCGCC----- 5015
898 GlyProAlaIleAlaIleAsnIleLysSerProProValIleuLys---SerIleAlaIleAla 917
5016 -----ACCAACAACA 5026
917 rProValHisIleuLysSerProSerIleuProAlaProSerProGlyYTrpThrSerSerPr 937
5027 TGACCCCAAACTTCGATCCCATCTCTTTCGACAAATGTCACCAACAATCAAGATT 5086
937 LysProProIleuGlnSer-----ProGlyIleProProAsnHisLysAla 952
5087 GCGTTCACGATGAGTATCCAGCCGGAACATGCTGTATAGCAAGATGAGGGGGTCCG- 5145
952 aProIleu-ThrMetAlaSerProAlaMetIleuLysAsnValGlnSerGlyIleProPro 972
```

```

QY 5146 -----GGACCGCTCGCATCATTAATTGAG----- 5169
Db 972 roProthralaserGlnProAlaserValAsnIleProGlySerLeuProSerSerThrp 992
QY 5169 ----- 5169
Db 992 roTyThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSera 1012
QY 5170 -----CATGAAATAATTTCCGGAC 5188
Db 1012 rgMetSerIlySpheAlaMetProSerSerThrProLeuTyThrIleAspAlaIleIleThrv 1032
QY 5189 TAGATGATCGAGATAGATPACC-----ATAAATCAAAATACTGTCAATTCATGATG 5242
Db 1032 aAlaserSerAspAspSerProProAlaIleArgSerProAsnLeuProSerMetAsn- 1051
QY 5243 TCGTATGAACTCAATGGGTCCCGGAATGCTGAATCTTAATATGTCGATGACAGGGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle-----SerGlyP 1068
QY 5303 CAAT----- 5307
Db 1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
QY 5308 -----GACCGCTGGCTTAATCTTAATTCCTCAATTCCTCAATTCCTCAATTCCT 5350
Db 1088 iAserAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProhIleGlyVal- 1107
QY 5351 GAGAGAATTCATAGAGGTCTGGTGGTGGCTGACAGCAACTTCAAACTTCAAGGGGTG 5410
Db 1108 -----ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
QY 5411 TTCACCTGGTGGCAGAAATGATGGTGAATGCAATGCAATGTTGGTGAATTTGAATC 5470
Db 1125 lngIuProProMetValProGlnGlyArgMet-----GlyPheProGlnGlyPheProP 1143
QY 5471 CGAATATTCAAGTAAAGGAGATACCCCAACCATACATACATACATCCAGTAAGGCAC 5530
Db 1143 ro-----ValGlnSerProGlnGlnValP 1152
QY 5531 AGAACGCCAACAACATTAACAACATGAGCTAATAATGCGAATGCCAATCGTACGTGG 5590
Db 1152 roPheProhIleAsnGlyProSerGlyGlyGlnGlySer-----PheProGlyGlyMetG 1170
QY 5591 AATTTTTGAGAGGTAGCTTAACCTCAATGGGTCTGTAGAGCAATGGGTCCCAATG 5650
Db 1170 lyphe-----ProGlyGlnGlyProLeuGlyArgProSerAsnLeu- 1183
QY 5651 GCCCACCATCAGCCAGCAGCGTACTCTCGAAATGCCAGATTGATGCCGAGCAGAG 5710
Db 1184 --ProGlnSerSerIleAspAlaIleAlaLeuCylylsProGlyGlyProGlyGlyProAsps 1203
QY 5711 CCGAGAGTATGCTAATGAATCT- 5733
Db 1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223
QY 5733 ----- 5733
Db 1223 allIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243
QY 5734 -----TCCGAGAGCAACACCGAACAAGATCA 5761
Db 1243 lulySerSerGlnThrLeuGlnTyPheProArgGlyGluValProGlyArgGlyGlnP 1263
QY 5762 CAACAATCTCTGGGGCAGCAATGTAATTACTTTTGAGATTTGCAATCAATGCTCA 5821
Db 1263 roGlnGlyProGlyPro-----GlyPheSerHisMetGln-----GlyW 1276
QY 5822 TTGTTGACGAGAAGGT-----GATTAACCCGCGCATGACGATCATCA 5866
Db 1276 eMetGlyGlnGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296
QY 5867 ATATTGTCACCATCTATGATTAAGGGGCAATGCGTCA-----CATGCCA 5911

```

```

Db 1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316
QY 5912 TGGCGCCCAAT-----GTAATGGGTGGCGGATGCACCCGCTTAACAGCAAAATTCAGT 5965
Db 1316 eArgProProAlaPheLeuGlnGlnGlyMetMetGlyProhIleHisArgMetMetSerP 1336
QY 5966 TTGCACAGTCATCGATGGTATTTGACTGTGCGGAGATCCGTCATCATTTTCACTAAG 6025
Db 1336 roAlaGlnSerThr-----MetProGlyGlnProThrLeuMetSerAsnProA 1352
QY 6026 CTTCCTGCAACAGCGCTGGACCAACATGTTTGATCAGACAAACAGCCATCAGCCTA 6085
Db 1352 lAlaAlaA----- 1354
QY 6086 AGACACAACAGCAATTAAGAAATACCTAGTGAATGTGCAAAACCAATCGGACTTGAC 6145
Db 1355 -----ValGlyMetIlePro--GlyLysAspArgGlyProAlaGlyLeuTyT 1370
QY 6146 TGGCAACAGGCGAGATCCCAACTGCATGGGCAAGACATGCGAGGTCAGTCTTAATG 6205
Db 1370 hrHisProGlyProValGlySerProGlyMetMetSerMet--GlnGlyMetMetG 1389
QY 6206 GACCTTAATAATAATTA-----ATGTCAACTGCCGGAAGTGCAGTCTCA 6253
Db 1389 lypProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly--MetAlaAlaA 1408
QY 6254 CTACGCGTGTCTGTGCAATTCGTAGTCCCTCTCT 6294
Db 1408 spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn 1420

RESULT 11
ADJ70152
ID ADJ70152 standard; protein; 1426 AA.
XX
AC ADJ70152;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target segid1958.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarrhythmic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200308768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-038987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
XX MPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function.
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.

```


Db	1152	roPheProHLeAsnGlyProSerLeItyGlyGlnGlySer-----PheProGlyGlyLeuMetG	1170
Qy	5591	AAATTTTGAGAGGTAAGCTTAACCTTCAMATGGGTGCTAGAGCAATGGATCGCCAAATAT	5650
Db	1170	lyPhe-----ProGlyGlnGlyProLeuGlyArgProSerAsnLeu-	1183
Qy	5551	GCCCAACCATTCAGCCAGCGAGCTACTCTGTGAATGCCAGAGATTGATGGCGGAGACCAAGAG	5710
Db	1184	--ProGlnSerSerLeAsnAlaLeuCyGlySerProGlyGlyProGlyProGlyProAsps	1203
Qy	5711	CCGAGAGATGCTATGAAATCT-----	5733
Db	1203	erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV	1223
Qy	5733	-----	5733
Db	1223	allLeArgProGlyAlaThrGlyLeProGluPheAspLeuSerArgIleIleProSerG	1243
Qy	5734	-----TCGGAGAGCAACACCAGAACAAATGCA	5761
Db	1243	IuLySerProSerGlnThrLeuGlnItyrPheProArgGlyGlyAlaValProGlyArgIySerGlnP	1263
Qy	5762	CAAPCAATCCGGGGGCAAGCAATGGTATTACTCTTTGAGAAATTGCAATGAATGTCTA	5821
Db	1263	roGlnGlyProGlyPro-----GlyPheSerHisMetGln-----GlyW	1276
Qy	5822	TTGTTGACAGAGGGGT-----GGATTACCCGGCCATGACGGATCAATGA	5866
Db	1276	etMetGlyGlnGlnAlaIleProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP	1296
Qy	5867	ATATTGTGCAACCATTTATGATTAAGGGGCAATGCGTCCA-----CATGCCA	5911
Db	1296	roValGlyThrProAspIleProLeuGlnItyrAlaProSerMetProGlyHisAsnProM	1316
Qy	5912	TGCGGGCCAAAT-----GTAATGGGTGGCGGAGATGCCACCCGTTAAACGAGCAATTCAGT	5965
Db	1316	etArgProProAlaIlePheLeuGlnGlnGlyMetMetGlyProHisIleArgMetMetSerP	1336
Qy	5966	TTGGACAGTCATCGGATGATTTGATCTGTGTCGGGGATCCGTCATCATTTTTCACATAACG	6025
Db	1336	roAlaGlnSerThr-----MetProGlyGlnProThrIleuMetSerAsnProA	1352
Qy	6026	CTTCCTCGCAACAGCGCTGGACCAACATGTTTGATCAGCAACAGGCCAATCAGCCTA	6085
Db	1352	IaAlaIaIa-----	1354
Qy	6086	AGACACCAACATATAAGCACTTACCTAGTGAATGTGTCAAAACCAATCGGACTTTGCAG	6145
Db	1355	-----ValGlyMetIlePro-----GlyIyAspArgGlyProIaGlyLeuItyrT	1370
Qy	6146	TGGCACAAGGGCAGATCCCACTGCATGGAGGGCAAGACATGCGACAGGCTACGCTTAAATG	6205
Db	1370	hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnIlyMetMetG	1389
Qy	6206	GACCTACTAATAATAATTA-----ATGTCAACTGCGGAAAGTGTCAAGTGCTA	6253
Db	1389	lyProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly---MetAlaIaIa	1408
Qy	6254	CTAAACGCTGTCTTGGCATCAATTTGTAAGTCCCTCTTCT	6294
Db	1408	spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn	1420

Human; legless; lgs; cell differentiation disorder;
KW cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;
KM breast; head; neck; brain; thyroid; skin; blood disease;
XX tissue regeneration; tissue repair; cytosstatic; lgs/Bcl9.
OS Homo sapiens.
XX
EN US2004038901-A1.
PD 26-FEB-2004.
PP 22-SEP-2003; 2003US-00664859.
PR 28-JUL-2000; 2000US-0221502P.
XX 27-JUN-2001; 2001US-00915543.
PA (UYZU-) UNIV ZURICH.
PI Basler K., Brunner B., Froesch B., Kramps T., Peter O;
DR WIJ.; 2004-203288/19.
XX DR N-PADB; ADJ71902.

Novel polypeptide sharing one or more homologue amino acid domains with Legless protein being functional homologue of Legless, useful for diagnosing disorders of cell fate.

Example 2; SEQ ID NO 15; 62pp; English.

The invention relates to a polypeptide sharing one or more homologous amino acid domains with a Legless (lgs) protein also related to a nucleotide sequence encoding a protein present in invertebrate and/or vertebrate organisms, the nucleotide sequence encoding a protein comprising a positive function in a regulatory pathway and the use of the polypeptide for the isolation of lgs-binding proteins by carrying out an assay chosen from an in vitro binding assay with such a peptide or a co-immunoprecipitation from vertebrate or invertebrate cell lysates or a mammalian or yeast two hybrid assay. The polypeptide and polynucleotide are useful for treating disorders of cell fate, which involves administering therapeutic compounds chosen from invertebrate and vertebrate lgs protein homologues or fragments, antibodies, antibody fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA, small peptides or chemical and natural compounds being capable of interfering with lgs function, synthesis and degradation. The disorders are related to cell differentiation or cell proliferation. The compound is administered to treat a cancerous condition by preventing progression from a pre-neoplastic or non-malignant condition to a neoplastic or malignant state. The cancerous condition is characterised by over-stimulation of the wnt pathway and is medulloblastoma or cancer of the colon, breast, head and neck, brain, thyroid or skin. The therapeutic compound may also be administered to a blood disease to promote tissue regeneration and repair. This sequence represents the human Lgs/Bcl9 polypeptide of the invention.

SQ Sequence 1426 AA:

Db 116 ---ProAsnAspSerSerPileuylGluCysAsnSerAlaAspHisIleuylSerGln 134
QY 2514 GAATCTTTAATATATCTCGCAATTTTCAGCATATTTAGTAAGCGAGTAGGACCAATACCC 2573
Db 135 AspSerGlnHisThrProHisSerMetThr----- 144
QY 2574 GGAATCGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCCATGGAAATCTCC 2633
Db 145 -----ProSerAsnAlaThrAlaProArgSerSerThrProSer 157
QY 2634 TCGCGTAGCAGTAATTTGTTGATTACATCCAAACACCAAAATAC-----ATA 2681
Db 158 HisGlyGlnThrThrAlaThrGlnProThrProAlaGlnThrThrProAlaIysValVal 177
QY 2682 TTGCTTTTTCATCTCAGCTGCGCAACAAAGCGCGCAATCAGTTTAAACGGCTCAATTT 2741
Db 178 TyrValPheSerThrGlnMetAlaAsnIlyAlaIleGlnAlaValIleuylSerGlnIleVal 197
QY 2742 CAAACTATTATTTCGGTATCATCTGCACCTCAGCCTGTCAAAAAAGCTTCTCGAAGACTTT 2801
Db 198 GlnThrIleValSerPheHis----- 204
QY 2802 TTTATGAAAAACCCCTTAAAGATTAAAGATTACAGCGGACAAATTCGGCTATGCCA 2861
Db 205 ---IleGlnAsn---IleSerAsnAsnIlyThrGlnArgSerThrAla-----Pro 219
QY 2862 TGGATAGGCGATGGGGGAGGTGGACTTAACCTCTCTAATCTGTAGCCAAATAACAA 2921
Db 220 LeuAsnThrGlnIleSerAlaIleuArgAsnAspProIylSerProleu-----Pro 235
QY 2922 CAGCAGCCACATACAAAGACCGTAGGCTATTGAAACCCCAATTCAATCAATGAAAC 2981
Db 236 GlnGlnProProAlaProAla-----AsnGlnAspGlnAsn 247
QY 2982 AGCAACGTAGTACTGTAAAGCGCGCTAGCACTCTTT----- 3020
Db 248 SerSerGlnAsnThrArgLeuGlnProThrProIleProAlaProAlaProIylSerPro 267
QY 3021 -----GTCCAGCAGTCTGATCTTATGGGCAACGAAACTGATGATG 3062
Db 268 AlaAlaProProArgProleuAspArgIylSerPro---GlyValGlnAsnIylSerIle 286
QY 3063 TGTGGGAAGCGGATCTCTCAACACCACT-----AGTCTGGAACAAAC 3107
Db 287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306
QY 3108 TCACGAAATCATGTACAGCATACGTACATCCAGAGGACACAGCCAAATAAAGATAC 3167
Db 307 SerThrProAsnAsnArgAlaValIThrProValSerGlnIylSerAsnSerSerAla 326
QY 3168 GAAGCAGCTGGCGGTGATTGGACAGGTTCACAAAGAAAGGAGATCTGGCTGACAACT 3227
Db 327 AspProIylSerAlaProProProProValSerSerGlyIylProProThrIleu---Gly 345
QY 3228 GAAAAACAACCTGTATCATCTGCAAGAGATTAAAGTTCCAGACGAAACCTTACACACA 3287
Db 346 GlnAsn-----ProAsp---GlyLeuSerGlnIyl 354
QY 3288 CACGGGCAATCGGGAAGAAGACAGTTGGCAAAAAATAAAAAATGAATCAATCTCTTTT 3347
Db 355 GlnLeuGlnHisArgGlnArgSerLeuGlnThrLeuArgAspIleGlnAsnMetLeuPhe 374
QY 3348 CCTGAAATATGAGATTCAGTAGAGCTAATGTAAAGCTCACAGATTAACAAAAAATTCCAGA 3407
Db 375 ProAspGlnIylSerGlnPheThrGlyAlaGln----- 384
QY 3408 GATTTAATGATGGGAGATGTCGGGTGGCGGAGCTATATTAATTAATCCAGAGATGCCA 3467
Db 385 -----SerGlyIylProGlnIlnAsnPro----- 392
QY 3468 CAACTGCATATGCCAGGTAAACGCCAAATCGGAGCTCTTATCGGCGACAACTTGACAGACT 3527
Db 392 ----- 392

QY 3528 TCGAAGATGTATGATCATCCAGGGATGTTATATCATGATATGGTGCCGTAATAGATGT 3587
Db 393 -----GlyValLeuAspGlyPro 398
QY 3588 AATTAATATCAAAAAACAGTGTGCATATGTGATCTCGAGTAGGTGTGTCTACTGGAA 3647
Db 399 GlnIylSerProGlnGlyProIleGlnAla----- 408
QY 3648 ACTGCAGCTGAGTAAATGTCAATATGCTCAATGCTCCAGCGCCCGCAATGGCAAT 3707
Db 409 -----MetMetAlaGlnSerGlnSerLeuGlyIylSerGly---- 419
QY 3708 ATGATGGGAAGCTTACCGATATGCTAGCTGTTGGCAACAAAGTGCACATCTCATC 3767
Db 420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnIylHisArgAspVal 438
QY 3768 GGAACGCCCCAGAT---ATGCTAAGGAAGTTTAAATCAAGATAGCCGAAC----- 3818
Db 439 ProPheSerProAspGlnMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458
QY 3819 ---CATTCACATCAAGGGGAGTGTCTCAATAGAGTGGTCGAAGATTCACATCA 3872
Db 459 ProAspHisLeuAspHisMetThrProGlnGlnIleAlaThrLeuIylSerLeuGlnGlnIyl 478
QY 3873 TTTTTCGAAGAACGCTCAAGGGGGGCAAGCCCAAGCAAGTCACTGGAATCTGATGACA 3932
Db 479 PheTyrGlnGlnIylSerArgArg-----LysGlnGlnIlnVal-----ValValGln 493
QY 3933 CAA-----CAGCAAAACCCCTTTCGGATCTGTGTGGAAC 3965
Db 494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgIyl----- 509
QY 3966 TCGTTAAACAACAGGTGGCAACCTCGCAAGGTCCACCTCTCTTACAC----- 4016
Db 510 -----ValValArgGlyProProProProIylProIylMetThrPro 522
QY 4017 -----TCCATCCAGAGATCTGCG 4034
Db 523 SerGlnGlyTyrAlaProGlyIylThrGlnProPheSerAspGlyIleAsnMetProHis 542
QY 4035 TCAGTACCAATAGCCACTCAATCCGCCCAATCCCTCGAGTCCAAACAATCTA----- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProIylSerIlnMetArgLeu 562
QY 4086 -----TCTCTCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGlnMetGlnIylProAsnValProAsnProAla 582
QY 4104 ACAACCGCAGCATGTGGGATG-----CGACCAACTCTTACGATG----- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTyrProAspAspValProIylSerIleProAspGly 602
QY 4149 -----GATGGACAGATCATTAATCTGCA----- 4172
Db 603 ArgAsnPheProProGlyIlnGlyIlePheSerGlyProGlyArgGlyIylArgPhePro 622
QY 4173 -----TCTGTTCCGCAAGCTTAATATCTGACGCGTTCAGCGAGCGC 4211
Db 623 AsnProGlnIylLeuSerGlnIylMetPheGlnGlnIlnLeuAlaGlyIylSerGlnIlnGly 642
QY 4212 ACAACA-----ACAGTCTCTCAGCAACAAAGACTGTTTTCAGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGlnIylIleArgProSerMetGlnMetAsnArgMetIle 662
QY 4266 CGTCAATATCAAAATGTGTAGTAAATACCGGATGTCAAGGTTCTTACGATTAATCTTA 4325
Db 663 ProGlySerGln---ArgHisMetGlnProGlyAsnAspProIlePheProArgIlePro 681
QY 4326 AGCAGCAACCCCAAGTACCCCTTATCTCATCTATCCCAAG----- 4367
Db 682 ValGlnGlyProLeuSerProSerArgIylAspPheProIylSerIleProProGlnMet 701

QY 4368 -----GAATTTGAGCTTTTCGTCAGTCTCTGCTGATGATATATTTGTT 4415
 Db 702 GlyProGlyArgGluLeuGlu-----PheGlyMetValProSerGlyMet----- 716
 QY 4416 AATTTTAAAGCAAAATCAATATCAATATGCGTTAATATAAGTTATATTACATAA 4475
 Db 717 -----LysGlyAspValAsnLeuAsn-----ValAsn----- 725
 QY 4476 CTCGGAATTTGATAGAAAAATCAGAAATAGAAAAATTAATTTATTTCCGACCG-- 4532
 Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGln 745
 QY 4532 ----- 4532
 Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
 QY 4533 -----CCCATCATTTCTTG 4547
 Db 766 ProLeuProPheGlyGluHisProGlnGlnGlyIuTyGlyMetGlyProArgProPheLeu 785
 QY 4548 -----AATCCAAATTTCTGAGTATTTGATAGATAA----- 4579
 Db 786 PrometSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
 QY 4580 -----TCTACTATTAAATTAACACGAAATTCATATCCGTTAATTGAAAT 4627
 Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProLeuProLeu----- 822
 QY 4628 CACTATTGTTTAAATTAAGAAATTAATAATATGTTATTAATATTCTACAGGTGATAC 4687
 Db 823 -----AsnProSerSerAsn 827
 QY 4688 ATGAAAGTAGGCGCAACCCACAGGTCAGCGTCACCAAGTAATAGTCTAATAGAG 4747
 Db 828 ProThSerLeuAsnThrAlaProProValGlnArg-----GlyLeuGlyArg 843
 QY 4748 GCAAAATAAGATGTACGATTTGCTGATCC-----AGTCGTGTTTAAACCGCAT 4798
 Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro-- 862
 QY 4799 CCACATATGCAAAATTCMAATTCAGATTAAACGCTATTAATAGGCTCATCAAT 4858
 Db 863 -----LeuLysSerProThMetHisGlnValGlnSerProMetLeuGlySerProSer 880
 QY 4859 ATACAGATGAGGTAATATTATTAATTTTATTAAAGTTTGTGTTAATTATCTTC 4918
 Db 881 GlyAsnLeu----- 883
 QY 4919 TTTTTCAGCGTCAGATCAGCCGCAAGGTGATCCGATCAATTAGTCGC----- 4969
 Db 884 -----LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
 QY 4970 -----GCTCCGATTAATATTCGCTTAATCCCAATAGTCGCAATCGCCGCGC----- 5015
 Db 898 GlyProAlaIleAlaIleAsnSerIleLysSerProProValLeuGly-SerIleAlaIle 917
 QY 5016 -----ACCAACACAGA 5026
 Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpHisSerSerP 937
 QY 5027 TGACCCAAACTCGATCCAAATCTCTTTGGCAACAAATGCCCAACAACATAAGT 5086
 Db 937 OlyAspProLeuGlnSer-----ProGlyIleProLeuAsnHisIlyVal 952
 QY 5087 GCGGTGTCAGCATGGTATGTCAGCCGGAACGTGGTGTATGACGATGAGGGGTCG- 5145
 Db 952 aspLeu-ThrMetAlaSerProAlaMetLeuGlnValGlnSerGlyGlyProProp 972
 QY 5146 -----GACCGTCGCACTCAATATTGAG----- 5169
 Db 972 robProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThp 992
 QY 5169 ----- 5169

Db 992 rofYrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
 QY 5170 -----CATGAAATATTTCCGGAC 5188
 Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032
 QY 5189 TAGATGATAGGAATAGATAC-----ATAAATCAAAATACTGTCATTCATGAATG 5242
 Db 1032 alaLaserSerAspAspSerProProAlaArgSerProAsnLeuProSerMetCys- 1051
 QY 5243 TCCTAATGAATCAATGGTCCCGAATGCTGAATCTTAATATGGTGTAGACGCGTC 5302
 Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle-----SerGlyP 1068
 QY 5303 CAAT----- 5307
 Db 1068 roAsnProValValPrometProThrLysSerPrometGlyMetThrGlnProLysSerH 1088
 QY 5308 -----GACCGCGCTGGCTTAATCTAATCCCAATGGGATTA 5350
 Db 1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
 QY 5351 GAGAGAAATTCATAGGAGTCTGCTGCTGCTCAGCAAACTTTCAAACTTTCAAGGGGTTG 5410
 Db 1108 -----PrometGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
 QY 5411 TTCACCTGCTGCCAAGATGATGGTGAATGCCAGTCAATTTGTTGTCGAATTTCAATC 5470
 Db 1125 InGluProPrometValProGlnGlyArgMet-----GlyPheProGlnGlyPheProP 1143
 QY 5471 CGAATATTACGTAAGGAGGAGTACCCCAACACCAATACATATACATCCAGTAAAGGCAC 5530
 Db 1143 ro-----ValGlnSerProProGlnGlnAlaP 1152
 QY 5531 AGAACCCCAACAATAACAAATGAGCTAATATGAGCAATGCCACTGATCTGCG 5590
 Db 1152 roPheProHisAsnGlyProSerGlyGlyGlnLysSer-----PheProGlyGlyMetG 1170
 QY 5591 AATTTTTCAGAGTACGCTAACCTTCMAATGGGTCTGTAGCAATGGTCCCAATAT 5650
 Db 1170 LysPhe-----ProGlyGlnLysProLeuGlyArgProSerAsnLeu- 1183
 QY 5651 GCCCACCATCAGCCACGACGAGCTCTCTGGAATGCGAGATTGATGCGGACGACGAG 5710
 Db 1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203
 QY 5711 CCGAGGATAGCTAATGAATTC----- 5733
 Db 1203 erPheThrValLeuGlnAsnSerMetProSerValPheThrAspProAspLeuGlnLuv 1223
 QY 5733 ----- 5733
 Db 1223 allIearProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243
 QY 5734 -----TCCGAGAGCAACACACAGAACATCA 5761
 Db 1243 IuLysProSerGlnThrLeuGlnTyPheProArgIleGlyValProGlyArgLysGlnP 1263
 QY 5762 CAACAATCTGGGCGCAGACGAATGATTAATCTTTTCGATGATGCAATCAATGCTCA 5821
 Db 1263 roGlnGlyProGlyPro-----GlyPheSerHisMetGln-----GlyM 1276
 QY 5822 TTGTTACGAAAGAGGT-----GGAATTAACCCGCGCATGACGATCAATGA 5866
 Db 1276 etMetGlyGlnGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296
 QY 5867 ATATTGTCACCATATGATAGGATGAGGATGCTCA-----CATGCCA 5911
 Db 1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316
 QY 5912 TCGCGCAAAAT-----GTAAATGGGTGCGCGAGTGCACCGGTTAAGGCAATTCAGT 5965

Db	426	-----	ProheuleuysglYProProglYglYAlaGly	436
QY	3444	TCATATTAATCCGACGATCGACAACTGATATGCGAGTAAACGCCAAATCGGAGCTC	3503	
Db	437	-----	-GluYglYProProAlaGlnAla	444
QY	3504	TTATCGCGCAAGTTCAGACTTTCCGAAGATGMAATGACATCCAGGGAGATTATATCA	3563	
Db	445	ProserAlaAlaGlnProProProserAlaPro	459	
QY	3564	GATATGGTGCCGTATAGATGTAAATATCAAAAACCAAGTGTCAATGTGTGATCT	3623	
Db	460	-----	-LysLysTyrGluGluProLeuGln	467
QY	3624	GGAATAGGTGTGTCACTGCAACAACTGCACTGGAGTAAATGTCAATATGCTATGCTCA	3683	
Db	468	-----	-SerMetIleSerGlnThrGlnSerLeuGluYProProLeuGlnHis	484
QY	3684	AGCTCCGGCGCCCGCAATGTCGAATATGATGGAAAGCTTACGATATGCTAGCTCGTTT	3743	
Db	485	ValProGluHisProGlnGluYAspMetGluGlnMetAspMetMet	501	
QY	3744	GGCAACAACACTGCAACTGATCGGAACGGCCCAAGATATGTCAAGAAATTTAAT	3803	
Db	502	-----	-MetGlnArgLeuLys	506
QY	3804	CAGAATAGCCGAACCCATTCATCAAGGGAGATTGCTCAATGAGATGGTTCGAAGATT	3863	
Db	507	GlnAspSerLeuThrProGlu	520	
QY	3864	CAACATCAATTTTTCGAAGAACGC	3899	
Db	521	GlnGluGluTyrTyrGluGlnLysArgLysGluGluGlnGlnLeuHisValGly	540	
QY	3900	AAGCCGACGAAGTCACTGGAATCTGTATACCAACAACGAACCCCTTCGAGATCGT	3958	
Db	541	ArgPro	550	
QY	3960	GGAAACTGTTAAACAACAGATGGCAACCCCTGCAAGGTTCACCCCTTCATCAACATCC	4019	
Db	551	Gly	563	
QY	4020	-----	-ATCCAGAGATCGGCTCACTCAATAGCCACT	4052
Db	564	LysProGluAspGlnCysAlaProGluAlaProGluAlaGlnLeuAspGluProMetAspVal	583	
QY	4053	CAATCGCCC	4103	
Db	584	GlnAspPrometGlnLeuAspProGluProPro	599	
QY	4104	ACAACCGGACAGTCATGGATTTGCCGACCAACTCT	4151	
Db	600	-----	-PheProLysGlnMetGlnAspValProGluPheGly	612
QY	4152	GGAACAGGATCATTTCTGATCTGTTCG	4208	
Db	613	GlyMetGlnSerMetPrometGluValProMetAspAlaMetGlnArgProValArgPro	632	
QY	4209	GGCACAACA	4253	
Db	633	GlyMetAlaTyrAsnGluAspLeuProProLysGluYProSerAspAspPheAlaGlnAsp	652	
QY	4254	GACACCCCAACCGCTCAATCAAAATGATGTAGATAATTACCCGATCGCAAGGCT	4310	
Db	653	AlaValProTyrProGluYglGlnLysAlaGluAlaGluAspGlyMetThrProArgValArg	672	
QY	4311	-----	-CTTAAGCATTAAGTAAAGCAACCCCAAGTACCCCTTATGATCATATGATCC	4361
Db	673	GluGluLeuLeuAspGlnGlnLeuLeuGluLysAspSerMetGluMetGln	689	
QY	4362	CCAAAGAAATTGAGCTTTGGTGCACAGCTCTGCTGGTATGTTATTTGTTAAATTTT	4421	

[illegible]

QY 5273 TGAATCTTAATAATGCTGAGGAGGCGCTCAATGACCGCTGCTT-----AATC 5326
 Db 993 yserProserMetAlaValProser-----ProGlyTTPvalAlaSerP 1008
 QY 5327 CTAATTC-----CCCAATGTGTGATTAAGAGAAAT-----TCCATAG 5365
 Db 1008 rolyThrAlaMetProserProGlyValSerGlnInuSerGlnProProLeuSerIleA 1028
 QY 5366 GGTCTGCTGTGTGCTGAGCAAACTTTCAACTTTCAGAGGCTGTTCACCTGAGCC- 5424
 Db 1028 snSerSerSerThrLeuGlyAenValGlu-----GlnGlyAlaLeuProProSerAlaP 1046
 QY 5425 ----AGATGATGGGTGCAATGCCAGTCAATTTGGTTCGAAT---TTCATCCGAATA 5476
 Db 1046 roArgAnSerSerSerAlaProProAlaAnProSerSerGlyLeuMetAnPro---- 1064
 QY 5477 TTCAGGTAAGGAGGAGTACCCCAACCCATACATATACATGCAAGGAGGACAGAGC 5536
 Db 1065 -----SerLeuProPhetThrSerSerProAerProThrProSerGlnAn- 1079
 QY 5537 CCAACAACATATACAACATGAGCTAATATGTCGAATGCCAGTACCTAGTCTGAATTTT 5596
 Db 1080 -----ProLeuSerLeuMetMetS 1086
 QY 5597 TGCAG-----AGGTACGCTTAACCTCAATGAGTGTGAGCAATGGGTG- 5643
 Db 1086 erGlnMetSerIyTyAlaMetProSerSerThrProLeuTyAlaSerAlaIleAla 1106
 QY 5644 -----CCAAATATGCCCAACCTATCAGCCA 5665
 Db 1106 hrIleAlaThrSerAerAerPgluleuProAerPglProLeuProProPro- 1125
 QY 5666 GCGACGCTACTCTGAGATGCCAGATGATGATGCGGAGCAGGAGCGGAGTATGCTAA 5725
 Db 1126 -----ProProGlnInGlySerGlyProGlyIleSerAn----- 1136
 QY 5726 TGAATTCCTCCGAGAGCAACACCAGAACAGATCAACAAACATCTCGGGGCAAGCAAT- 5784
 Db 1137 --AenGlnProAenGlnMetIleMet-----AenProAlaAlaIleGlnS 1151
 QY 5785 -----GGTATTAATCTTTTTCAGAAATTCGAATCAATGTCTATTTGAGAGAGAG 5836
 Db 1151 erProMetGlyMetAn----- 1156
 QY 5837 GTGATTTACCGGCGCATGACGATCATATATGTCACCATCTATGATAGAGGCA 5896
 Db 1157 ----LeuProGlyGlnGlnProLeuSerIleGluProProProThrMetLeuProSerP 1175
 QY 5897 TGGGTCCACATGCC-----ATGCGGCCAAATGTAATGGGTGCGCGATGCCAC 5944
 Db 1175 roThrProLeuGlySerAnIleProLeuHisProAenAlaGlnGlyThrGlyIleSers 1195
 QY 5945 CCGTTTAACAGGCAAAATTCATTTGACAGATCATCGATGCTATGACTGTGTGGGGATC 6004
 Db 1195 erGlnAnSerMetMet-----AlaProGlyGlyP 1206
 QY 6005 CGTCATCATTTTCACTACGCTTCTGCAACAGCGCTGAGCACACATGTTGGATCAG 6064
 Db 1206 roAerSerLeu-----AenAlaProCys-----GlyProValProSerSerSerc 1221
 QY 6065 CACAACAGGCCAATCAGCTTACAGACACAC----- 6096
 Db 1221 InMetMetSerPheProProArgLeuGlnInProHisGlyAlaMetAlaProThrGlyA 1241
 QY 6097 -----ATAAAGAACATACCTAGTGAATGTGTCAAAACCATCGGAGC 6139
 Db 1241 laGlyGlyProGlyLeuGlnGlnHisTyProSerGlyMetAlaLeuProGlnAerP 1261
 QY 6140 TTGCAGTGGCAAGGAGGAGATCCAGTGCATGGGCAAGGATGCGAGGTCACTTT 6199
 Db 1261 eurPro---ThrGlnProProGlyProIleProProGlnGlnHisLeuMetCyllySerGlyM 1280

QY 6200 TAATTGACCTTAATAATAATTAATTAATGTCACCTCCGGAAGTGTCACTACTACG 6259
 Db 1280 etThrGly-----ArgMetGlyAerAlaTyProProGlyValLeuProG 1295
 QY 6260 GTGCTCTGGGATC-----AATTGCTAGGTCCTCTTCTACG- 6297
 Db 1295 lYValAlaSerValLeuAnAerProGluLeuSerGluValIleArgProThrProThG 1315
 QY 6298 -----GACCTGAAGTAT-----GCCAGCAATATCACTAGTTTTCACG 6334
 Db 1315 lYIleProGluPheAerPheLeuSerArgIleIleProSerGluYerProSerSerThrLeuG 1335
 QY 6335 AGCAGTTATATGCTACCAACACAGAGTCAG-----CAACAACAGCATATGC 6382
 Db 1335 In-----TyRheProIySerGluAenGlnProProIyAlaGlnProProAenLeuH 1353
 QY 6383 ACCAGAGCAGCAGAGCAGATGATA-----ACAAATGCCCGCAATTTAT 6427
 Db 1353 lAerMetCAnLeuGlnMetMetMetAlaGluGlnThrProSerArgProProAenLeu- 1372
 QY 6428 CACCAATTCACGTTCTTTGTCAACAAATAAATTTTTCGCGCCT 6481
 Db 1373 --ProGly-GlnGlnGlyValGlnArgGlyLeuSerMetSerMetCysHisPro 1389
 RESULT 14
 AAE36095
 ID AAE36095 standard; protein; 1596 AA.
 AC AAE36095;
 DT 26-JUN-2003 (first entry)
 DE Drosophila sp. Mastermind (Mam) protein.
 KW Notch signal transduction; Mastermind; Mam; Mip1; Mip30; Mipe; cancer;
 KM gene therapy; fruitfly; cyostatic.
 OS Drosophila sp.
 PN WO2002102987-A2.
 XX 27-DEC-2002.
 PD 18-JUN-2002; 2002WO-US019189.
 PF 18-JUN-2001; 2001US-0299112P.
 PR (GEHO) GEN HOSPITAL CORP.
 PA Artavanis-Tsakonas S, Lake RJ;
 PI WPI; 2003-167509/16.
 DR N-PSDB; AAD54620.
 XX Inhibiting or agonizing Notch signal transduction in a cell for preparing
 PT a composition for treating or preventing cancer comprises contacting the
 PT cell with an antagonist or agonist of stimulation.
 PS Disclosure; Fig 1; 84pp; English.
 CC The invention relates to a method for modulating Notch signal
 CC transduction in a cell which comprises contacting the cell with an
 CC antagonist or agonist of stimulation. The invention is based on the
 CC interactions of Mastermind (Mam) protein with the Mip1, Mip30 and Mip6
 CC proteins. The method is useful for preparing a composition for treating
 CC or preventing a disease e.g. cancer. The invention is also used in gene
 CC therapy. The present sequence is Drosophila sp. Mam protein used to
 CC illustrate the method of the invention
 XX
 SO Sequence 1596 AA;
 Alignment Scores: 1.84e-13 Length: 1596
 Pred. No.:

Score:	270.50	Matches:	239
Percent Similarity:	29.56%	Conservative:	148
Best Local Similarity:	19.78%	Mismatches:	544
Query Match:	2.25%	Indels:	521
DB:	6	Gaps:	70

US-10-664-859-1 (1-6909) X AAE36095 (1-1596)

Qy	2658	TACATCGAAGCAGCAAAATTCACATATTCGTGTTTCACTGCAGCTGGCCACAAAGAGGGCC	2717
Db	203	Hi e l n g i n g i n g i n h s -----G n g i h i s g i n g i n g i n h a	218
Qy	2718	GAATCAGATTAAAGCGGTCAATTTCAACTATTAATTGGTATCACTGCAGCTCAGCTGGCT	2777
Db	219	G l n h r h e t h e u a l a g l y G l n l e u g i n s e r v a l -----H s v a l G l n l y p h e	236
Qy	2778	ACAAAAGCTTCCCTGGAGACTTTT	2810
Db	237	Leu y s a r g P ro a l a g l a s p v a A s p a n g l P ro a s p e r h e g u P ro h i l y s	256
Qy	2811	AACCCCTTAAAGATTAAACAGTTTACAGCGGACAAATTCCTCGGTATGTCATGATGATAGC	2870
Db	257	Leu P ro a s p h e n s a n a s p h e n s e r v a s e r h e n s p h e n s a n g l y a s h a l a s h a a s h	276
Qy	2871	ATGGGCGACGTTGGACTTAATCTCTTAATCTGTGA	2906
Db	277	A s h l y l a n g l y s e r v e n h t h r g l y A s h a n h r a n h e n a n g l y a n s e r h t h r a n	296
Qy	2907	-----GCCAAATA	2915
Db	297	A s h a n g l y s e r v e n a s p h e n g l y s e r g l u b e n l e h r l y s h e s e r v a l u l e	316
Qy	2916	ACAAACAGCAGCAGCACATATACAAACCGTACGGCTTATGAAACCCCAATTCATATCAAT	2975
Db	317	V a l G l n h l e u e u g i u h e h r h h s e r P ro a l a s p h e r g i n P ro g i n -----	332
Qy	2976	GA AAAACAGCAAAACGTAGTACTGTAAAGCGCGCTTACCACTCTTTTGTGCAGACTGTAT	3035
Db	333	G l n h e r h t h r a n v a l h r v a l y a l e h r h a n h r s e r v a l -----L y s e r c l u	351
Qy	3036	CCT-----ATGGCAACGAAATCGAA	3056
Db	352	P ro g i l y a l y g l y g l y g l y g l y g l y a n s e r g l y a s h a n s p h e n -----	370
Qy	3057	TTGATGTCTGGGAAGCGGATCTCTCAAAACACAGTAGTGATGGACAAATCTCAGCAAT	3116
Db	371	-----A s h a n g l y g l y g l y g l y g l y g l y a n g l y a s h a n s h a n s h n	386
Qy	3117	CATGTAGACAGTATCAGTACATCCAGCGAGTCAACGAGCAATTAAGATCTAGGAACAGCT	3176
Db	387	G l y l y a s h i s h g i n g i n g i n h i e g i h i s g i n g i n g i n g i n g i n	406
Qy	3177	GGCGTTGATTTGGGACAGCTC-----ACAAAAGAGACGATCTCTGCG-----CTG	3221
Db	407	G l y l y g l y e u g i y g l y e u g i y a s h a n g l y a r g l y g l y P ro g i l y e h e r a	426
Qy	3222	ACAACGAAAAACAACATTGTA-----TCACTCGAAGAGTAAAGTTTCCGAGCAAAAACCTT	3278
Db	427	T h r g l y P ro g i l y e h a l a g l y l y e u g i y e h c l y e h P ro P ro a h e h e h	446
Qy	3279	ACACCAACACAGCGGCAA-----CATCGGAGAAACAGTTGGCAAAAATAAAAAAATG	3332
Db	447	S e r l l g i n g i n y s e r a h a l e u g i y a s h e n l a a s h e n v a h l u c l y a r g l u	466
Qy	3333	AATCAATTTCTTTTCTCTGAA-----AATGAGATTTCAAGTAGAGCTAATGTA	3380
Db	467	P ro a h i s a p h e P ro a h P ro a h e u g i y s e r l e h a s h y s a p h g l y g l y -----	484
Qy	3381	AGCTACAGATTAACAAAATTTCCAGGA-----GATTTAATGATGGGATGTCCGGTGGC	3434
Db	485	-----G n h e P ro g i l P h e P ro a h P ro a h e u -----G l y	494
Qy	3435	GGAAGCGGATCTATTATTAATTCGAGATGCGAACATCG-----	3473

[illegible]

[illegible]

QY	5439	AATCCAGTCGAATTTCGGT-----	TCGAATT	5465
Db	1029	rglnsermetchenphenenlnglnnglnnglnlnalalalaelnglnnglnnglnva		1049
QY	5466	CAATCCGAATATT-----	CAGTAAAGCGAGTACCCCAACCAACCATACATACAT	5516
Db	1049	lglnproabsmetearnglnarnglnthrglnlnalaglnlalalalalalalalalal		1069
QY	5517	GCCAGTAAAGGCGACAGAAAGCCCAACCAATTAACAAACAAATGAGCTAATATGCGCAAT		5576
Db	1069	aalalalaglnlalaglnlalalalalalalalasergly-----		1081
QY	5577	GCCACCTAGTCTGGAAATTTTTCAGAGGTACGGTAAACCTCAATGGGT-----		5625
Db	1082	-----ProabsnvalProleumetglnngln-----	ProglnvalglyalglyValgl	1097
QY	5626	-----	GCTGTAAGCAATGGGTCCGCATATATGCCACCATCAGC	5663
Db	1097	valglvalglvalglvalglvalglvalglvalglvalglvalglvalglvalglvalgl		1117
QY	5664	CAGCGAGCGTACTCCTGGAATGCGAGATGATGATGGCGGACCA-----	GAGCGCGAGG	5717
Db	1117	rgllyglProabsnannglnalmetaelnglnmetgllyglPrometgllyglmetProgl		1137
QY	5718	TATGCTATGTAATTCTTCCGAGAGCAACACAGAAAGATCAACAACATCTTGCGGGC		5777
Db	1137	ymetglnmetgllyglPro-----	metAsnPrometglnmetAsnProabnAl	1153
QY	5778	AAGCAATGATTTTAATCTTTTCAGAAATGCAATCAATGCTATTGTTGACGAAGAGGG		5837
Db	1153	alagly-----	Prothralaglnnglnmetethetglserglyalagl	1168
QY	5838	TGGA-----	TTACCCGCGCATGACGATCAATGAT-----	AT 5870
Db	1168	ylglProgllyglvalglProgllyProgllynglnlyProabsnproabnglnlalalysphete		1188
QY	5871	TGTCGAACCATCTATGATAGGGGCGATGCTCCACATGCCATGGCGGCAATGTAAATGGG		5930
Db	1188	uglnnglnnglnnglnmetmeteargalaglnalmetglnnglnnglnnglnnlnmetsergl		1208
QY	5931	TGGCGGATGCGCACCC-----	GTTAAACAGCGCAATTCAGTTTGACACATC	5975
Db	1208	yalalarProProProProglulutyraabnalathlyalalaglnneumetglnalaglnme		1228
QY	5976	ATCGGAT-----		GG 5984
Db	1228	tmecglnnglnthralvalglyglyglyvalglvalglvalglvalglvalglvalgl		1248
QY	5985	TATGACGTGTGGGGGATCGGTCA-----	TCATTTTCACTAAGCTTCTCCCAACAGCGC	6041
Db	1248	valglglyvalglglyalalabnglnlyglarPheProabsemearlalaglnlalal		1268
QY	6042	T-----	GGACCAACATG-----	6054
Db	1268	aalameetarargmetthrglnnglnProcoleProProbserclyPrometmetarPprogl		1288
QY	6054	-----		6054
Db	1288	nhlslametyrmetglnnglnlnlsglylvalgllygllyglProarhtrhlymetgl		1308
QY	6055	-----	TTTGGATCAGACACAAACAGGCGCAATCAGCC	6083
Db	1308	ylalProlyrtyrlygllyglarnglygllyglPrometcdlyglProodnglnnglnnglnnargr		1328
QY	6084	TAAACACAAACATA-----	AAGAACATACCTAAGGAAATGGTCAAAACCAATCGGG	6137
Db	1328	oProabnvalglvalThlProarPgllymetPrometcdlyserglnnglnnglnutrarPgrhl		1348
QY	6138	ACTTGCAATGGCACAAAGGCGAGATCCAACTG-----	CATGGGCAAGGACATCG	6185
Db	1348	smetethetthrglnnglnnglnthrglnmetgclPhegllygllyProgllyProgllygllyPr		1366

	RESULT 15	
XX	ABB63688	
ID	ABB63688 standard; protein; 1594 AA.	
AC	ABB63688;	
DT	26-MAR-2002 (first entry)	
XX		
DE	Drosophila melanogaster polypeptide SEQ ID NO 17856.	
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KM	pharmaceutical.	
OS	Drosophila melanogaster.	
NN	WO200171042-A2.	
XX		
PD	27-SEP-2001.	
XX		
PF	23-MAR-2001; 2001WO-US009231.	
PR	23-MAR-2000; 2000US-0191637P.	
PR	11-JUL-2000; 2000US-00614150.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
PI	Venter JC, Adams M, Li FWD, Myers EW;	
XX		
DR	WPI; 2001-656860/75.	
DR	N-PSTD; ABL07791.	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.	
PT		
PS	Disclosure; SEQ ID NO 17856; 21pp + Sequence Listing; English.	
XX		
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the published specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	
CC		
CC		
CC		
SO	Sequence 1594 AA;	
Alignment Scores:		
ired. No.:	2,24e-13	Length: 1594
score:	269.50	Matches: 295

[illegible]

Db 493 AasnSerGluAasnAspThrPheIysAspLeuIleAasnAsnLeuGlnAspPheAsnPro 512
 QY 3474 ---CATATGCGAGGTAAAGCGCAATCGAGCTCTTATCG---GCGCAAGATTGAGAGCTT 3527
 Db 513 SerPheLeuAspGlyPheAspGlyIysProLeuLeuAspIleIysThrGluAspGlyIle 532
 QY 3528 TCGGAAGATGTATATGATCGAGGGAGTGTATATCATGATATGAGTCCCGTAATAGATGT 3587
 Db 533 LysValGluProProAsnIleGlnAspLeuIleAsnSerLeu----- 546
 QY 3588 AATTAATATCAAAAAACCATGTGCAATGTGATCTGAGTAGTGTGTCTC----- 3638
 Db 547 ---AsnValIysSerGluGlyIleGluGlyIleGlyPheGlyIlePheGlyLeuGly 564
 QY 3639 ---ACTGAAACAATCGAGCTGAGTAAATGTCAATATGATGATGCTCAAGCTCC 3689
 Db 565 LeuAspAsnProGlyMetIysMetArgGlyIysAsnProGlyAsnGlnGlyIysPhePro 584
 QY 3690 GCGGCGCCGAAATGCAATATGATGGAAAGCTTACGAGATATGCTAGCCTGTTGGCAAC 3749
 Db 585 AsnGlyProAsnGlyIysThrGlyIysAlaProAsn-----AlaGlyIysAsnGlyIys 602
 QY 3750 ACAAGTGCACAGCTCATCGAAGCGCCCGATATGTCTTAAGGAAGTTTAAATCAAGAT 3809
 Db 603 AsnSerGlyAsnLeuMetSerGluIleAspPro--LeuAlaIleGlnThrLeuIysGlnMet 621
 QY 3810 AGCGCAACCATTCATCATCAAGGGGAGTGTCAATATGAGTGTGTCGAAGATTCAACAT 3869
 Db 622 AlaGluGlnIleAsnIleIysValAsnAlaMetGlyIysMet----- 634
 QY 3870 CAATTTTTCAGAGACGCTCAAGGGGGGAGGAGTGTCAATATGAGTGTGTCGAAGATTCA 3929
 Db 635 ---GlyGlyPheProArgProProIleGly--MetAsn 645
 QY 3930 CGACAAGACGAACCCCTTGTGATCTGTGTGAAACCTCTTAACAACAACGAGTGCAGACC 3989
 Db 646 ProGlnGlnGlnGln-----GlnGlnGlnGlnGln 656
 QY 3990 CTGCAAGTCCACCTCTCTTACCATTCATCCAGATCTGCTGACGACCAATAGCC 4049
 Db 657 GlnGlnGlnGlnIleGlnGlnGlnIleGlyIleMetMetGlyIysGlnProGlyIysArg 676
 QY 4050 ACTCAATCGCCCAATCCCTGAGTCCAAACAATCTATCTCTC-----CCGTACCGCGG 4103
 Db 677 TyrAsnAspTyrGlyIysGlyPheProAsnAspPheGlyLeuGlyProAsnGlyProGln 696
 QY 4104 ACAACCGGACATTCATGAGATTGCGGACCAACTCTCTAGCATGATGGAAGAAGATCA 4163
 Db 697 -GlnGlnGlnGlnIleGlnGlnGlnPro-----GlnGlnGlnIle 710
 QY 4164 T----- 4164
 Db 710 sLeuProProGlnPheIleGlnGlnIysGlyProGlyIysAlaGlyMetAsnValGly 730
 QY 4165 ---TATCTGATGTGTTCCGCAAGCTAATA-----CT 4193
 Db 730 ngIleAsnPheLeuAspIleIysGlnIleuPheTyrSerSerGlnAsnAspPheAspLe 750
 QY 4194 TCAGCGGTTCAGGACGACCAACAACAGTCTCTCAAGCAACAAGACTTTTTCAGGCA 4253
 Db 750 ulysAlaGlyLeuGlnGlnIleGlnIleAlaMetGlnGlnGln-----GlnGln 765
 QY 4254 GACAACCCATCGCGGTCAAAATCAAAATGTATAGAAATACCGGATGTCAGCGTCTT 4313
 Db 765 ngIleIleIleGlnGlnGln-----GlnGln----- 773
 QY 4314 ACCGATTAATTAAGACAGCAACCAAGTACCCCTTATCTCATCTTATCCCAAGAAATTT 4373
 Db 774 ---GlnGlnGlnProIys-----MetGlyIysValPro----- 783
 QY 4374 GAATCTTTGCGTCAAGTCTCTGCTGATATGTATATTTTAAATTTTAAAGACAA 4433
 Db 783 ----- 783

QY 4434 TCAAAATATGAATGGCTTAATATAGTATATATACATCACTCGAAATTTGATAGAA 4493
 Db 784 -----AsnPheAsnIysG 788
 QY 4494 AAAATCAGAAATAGAAAAATTAATTTATTTCCGAGCCGCCATTCATTTCTGAATCCA 4553
 Db 788 IlnGlnGlnGln----- 791
 QY 4554 ATTTGCGAGTATGTTAGATATATCTATTAAATTAACAGAAAAATTCATATTC 4613
 Db 791 ----- 791
 QY 4614 CGTTAATTGAAATATCATTTGTTAATAGAAATTAATAATGTTTATTAATATTT 4673
 Db 791 ----- 791
 QY 4674 CTACAGGTATATAATGAAAAATAGGCGACCAAGCCCAAGGTCAGCGGTCAACAGTAA 4733
 Db 792 -----GlnGlnValProGlnGlnIleGlnGlnGln 803
 QY 4734 ATAGTCTAATAGAGCAATAAAGATGTACGATTTGCTCATCTCAGTCTCTGTTTAAAC 4793
 Db 803 IlnGlnGlnGlnGlnGlnGlnGlnGlnGlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 823
 QY 4794 CGCATCCACATATGCAAGCAATTCAAATTCAGCATTAACGCTTAATAATGCGCTCTA 4853
 Db 823 IlnAsnProAsnIleAlaIleAsn-----PheLeuAsnIysProProArgGlyIys 840
 QY 4854 CCAATAT---ACAGATGAGGTAAATATTTAATTTTATTTAATCGTTTGTGTAA 4909
 Db 840 roAsnIleValAsnGlnGlnProGlyIys----- 848
 QY 4910 TTTATCTCTTTTTCAGCGTCAAGCATCAGCCAGGTGATCCGACATTTAGTGGC 4969
 Db 849 ---LeuAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnIlnIlnIlnIlnIlnIlnIln 866
 QY 4970 GC-----TCCGATTAATATTCGCTTAATCCCAATAGT-----GCGAATCGGC 5011
 Db 866 rgGlyIysAlaIleAlaAsnGlyIlnGlnIlnAsnAsnProAsnIleGlyProGlyIysAsnThr 886
 QY 5012 CGCCACCAACAAGATGACCCCAAACTTGATCCAACTCTCTTTGGCAACAATGCC 5071
 Db 886 roAsnIleProGlnGlnGlnGlnGlnGlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 905
 QY 5072 AACCACTAACAGTTCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 5131
 Db 905 IlnIlnGlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIlnIln 923
 QY 5132 TGATGGGGGGTCCGGGACCGTCCGACATCAATATGAGCATGAATTAATTTCCGAGCTAG 5191
 Db 923 alSerAlaGly-----GlnIleIleuIleIleuSerGlyIysAsp 935
 QY 5192 ATGATTCAGGAATAGTATACATTAATCAAAATATCTGTCAATCAATGATGCTGATATGA 5251
 Db 935 eelIysSerAsnValSerValAlaIleGlnGlnIln-----ValPhe 949
 QY 5252 ACTCAATGGGTCCCGAATGCTGAATCTTAAATGTGCGTACGAGCGGTCCAAATGAC 5311
 Db 949 heserGlnGlnGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 969
 QY 5312 CG---CTGAGCTTAACTCAATTTCCCAATGTGATTAAGAGAAATTCATAGGCT 5368
 Db 969 roAsnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 989
 QY 5369 CTGGCTGTGCG----- 5379
 Db 989 alGlyIysAlaGlyIysAlaGlyIysAlaGlyIysAlaGlyIysAlaGlyIysAlaGlyIys 1009
 QY 5380 ---TCAGCAACCTTCAAACTTTCAAGGGGTGTTCACCTGGTGCAGATGAGGCT 5437
 Db 1009 roAsnGlnIlnAsnMetSerAsn---AlaAsnValProSerAspGlyIysPheSerLeuSerGln 1028

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 22:24:08 ; Search time 1054 Seconds

(without alignments)
6713.387 Million cell updates/sec

Title: US-10-664-859-1

Perfect score: 12037
Sequence: 1 acgagtcgtctctctatctat.....aaaaaaaaaaaaaaaaaaaaa 6909

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 segs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q=/cg92_1/USPRO.spool_p/US10664859/runat_02112005_171807_3627/app_query.fasta_1.7047
-DB=uniprot -OFMT=faetan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=psco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10664859 -CCGN_1_1_1295 -@runat_02112005_171807_3627 -NCPu=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7280	60.5	1469	1 BCL9_DROME	Q961d9 drosophila
2	338.5	2.8	1530	2 O67FY3	O67FY3 brachydanio
3	322.5	2.7	1723	2 O8T248	O8T248 dictyosteli
4	320	2.7	1426	1 BCL9_HUMAN	O00512 homo sapien
5	313	2.6	1474	2 O67FY0	O67FY0 brachydanio
6	301	2.5	2232	2 O8IFX6	O8IFX6 caenorhabdi
7	297.5	2.5	1457	2 O641I9	O641I9 mus musculi
8	296.5	2.5	1494	2 O67FY2	O67FY2 mus musculi
9	289.5	2.4	1425	2 O67FY5	O67FY5 mus musculi
10	288	2.4	1462	2 O67FY9	O67FY9 mus musculi
11	281	2.3	1494	2 O81IR9	O81IR9 plasmodium
12	280	2.3	1499	2 O67FY1	O67FY1 homo sapien
13	279	2.3	1774	2 O81AU8	O81AU8 homo sapien
14	277.5	2.3	1902	2 O640Q1	O640Q1 plasmodium
15	277.5	2.3	3006	2 Q7YYS6	Q7YYS6 mus musculi
16	277.5	2.3	3006	2 Q7YYS6	Q7YYS6 cryptospori

17	274.5	2.3	1918	2 O86AF5	O86AF5 dictyosteli
18	270.5	2.2	1596	1 MAM_DROME	P21519 drosophila
19	269.5	2.2	1594	2 O9V6W7	O9V6W7 drosophila
20	267.5	2.2	1902	2 O92SQ1	O92SQ1 mus musculi
21	261.5	2.2	2206	2 Q7QAV8	Q7QAV8 anopheles g
22	261.5	2.2	2235	2 O81B09	O81B09 plasmodium
23	258.5	2.1	1655	2 O24754	O24754 drosophila
24	258.5	2.1	2151	2 O9VPL6	O9VPL6 drosophila
25	256.5	2.1	5332	2 O9VPL9	O9VPL9 drosophila
26	255.5	2.1	2151	2 O9Y1L3	O9Y1L3 drosophila
27	254.5	2.1	1348	2 O86KE2	O86KE2 dictyosteli
28	254.5	2.1	2038	2 O9W3L3	O9W3L3 drosophila
29	251.5	2.1	1502	2 O81S10	O81S10 dictyosteli
30	249.5	2.1	2038	1 FSH_DROME	P13709 drosophila
31	249	2.1	1685	2 Q7REY7	Q7REY7 plasmodium
32	248	2.1	3443	2 O8JZM8	O8JZM8 mus musculi
33	247.5	2.1	1902	1 SMF1_HUMAN	O14457 homo sapien
34	247.5	2.1	1999	2 O8NFD6	O8NFD6 homo sapien
35	247	2.1	1255	2 O9W3Q5	O9W3Q5 drosophila
36	246.5	2.0	2634	2 O63KC6	O63KC6 burkholderi
37	244.5	2.0	2072	2 Q7RSM8	Q7RSM8 plasmodium
38	243.5	2.0	1208	2 O6RTA9	O6RTA9 homo sapien
39	243.5	2.0	1323	2 Q7KS20	Q7KS20 drosophila
40	243.5	2.0	1374	2 Q7YU77	Q7YU77 drosophila
41	243.5	2.0	1376	2 Q9VI63	Q9VI63 drosophila
42	242	2.0	1353	2 Q7RDU8	Q7RDU8 plasmodium
43	241	2.0	796	2 O6NRE2	O6NRE2 xenopus lae
44	241	2.0	1608	2 O63UB0	O63UB0 burkholderi
45	241	2.0	2280	2 Q9VBB6	Q9VBB6 drosophila

ALIGNMENTS

RESULT 1	BCL9_DROME	STANDARD;	PRT;	1469 AA.
AC	O961d9; O9VAD2;			
ID	BCL9_DROME			
DT	28-FEB-2003 (Ref. 41, Last created)			
DT	28-FEB-2003 (Ref. 41, Last annotation update)			
DT	25-JAN-2005 (Ref. 46, Last annotation update)			
DE	Bcl-9 homolog (legless protein).			
GN	Name=1gs; Synonyms=BCL9; ORFNames=CG2041;			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkley;			
RX	PubMed=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	Georgie R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Mortman J.R., Vandeil M.D., Zhang O., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,			
RA	Abdel J.F., Agbayani A., An H.-U., Andrews-Piankocch C., Baldwin D.,			
RA	Ballew R.M., Baas A., Baxendale J., Bayraktaroglu L., Beasley B.M.,			
RA	Beeason K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Borchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke Z., Davenport L.B., Davies P.,			
RA	de Pablo B., Delcher A., Deng C., Maye A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,			
RA	Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hastati D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jostali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,			

RA Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moberly B., Murphy L., Murthy J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy J., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Weisenbach J.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [2]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Beutencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.U., Shu S.Q., Wang
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RL systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Garin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celisner S.E.;
 RL "A *Drosophila* full-length cDNA resource.";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 [4]
 RP SEQUENCE OF 6-1469 FROM N.A., AND MUTAGENESIS OF GLY-514; LEU-534 AND
 RP ILE-537
 RX MEDLINE=21952490; PubMed=11955446; DOI=10.1016/S0092-8674(02)00679-7;
 RA Kramps T., Peter O., Brunner E., Nellen D., Froesch B., Chatterjee S.,
 RA Murore M., Zuellich S., Baeler K.;
 RL "Wnt/wingless signaling requires BCL9/Legless-mediated recruitment of
 RT pygopus to the nuclear beta-catenin-Tcf complex.";
 RL Cell 109:47-60(2002).
 CC -1- FUNCTION: Involved in signal transduction through the Wnt pathway.
 CC -1- SUBUNIT: Binds to ARM and PYGO.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically
 CC throughout development.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 CC EMBL; AEO03844; AAF59345.2; -
 CC EMBL; AY051651; AAK93075.1; -
 CC DR EMBL; AF457205; AAL91368.1; -
 CC DR FlyBase; FBgn0039907; lgs.
 CC DR GO; GO:0005634; C:nucleus; IDA.
 CC DR GO; GO:0030528; F:transcription regulator activity; IPI.
 CC DR GO; GO:0030177; P:positive regulation of Wnt receptor signal.; IPI.
 CC DR GO; GO:0007067; P:segment polarity determination; IMP.
 CC KM Developmental protein; Nuclear protein; Segmentation polarity protein;
 CC Wnt signaling pathway.
 FT DOMAIN 511 555 ARM-binding.
 FT DOMAIN 1134 1173 Arm-rich.

FT	DOMAIN	1340	1449	Gln-rich.
FT	DOMAIN	1162	1166	Poly-Asn.
FT	MUTAGEN	514	514	G->E: in allele lgs-21L.
FT	MUTAGEN	534	534	L->F: in allele lgs-17E; segment polarity
FT	MUTAGEN	537	537	phenotype.
FT	MUTAGEN	537	537	I->K: in allele lgs-17P.
FT	SEQUENCE	1469 AA;	153759 MW;	5672B01B7200ED08 CRC64;

Alignment Scores:

Pred. No.:	0	Length:	1469
Score:	7280.00	Matches:	1465
Percent Similarity:	73.08%	Conservative:	1
Best Local Similarity:	73.03%	Mismatches:	3
Query Match:	60.48%	Indels:	537
DB:	1	Gaps:	5

US-10-664-859-1 (1-6909) x BCL9_DROME (1-1469)

QY	453	ATGCTCTGACAAACATGCGCCGAGCTCCACCCAAACACCGCAACCAATCCGANT	512
DB	1	MetLeuSerThrThrMetProArgSerProTrpGlnGlnProGlnProAsnSerAap	20
QY	513	GGCTCCTCAACAGGTCATCTGATCAATCTCGAGCAGCGAGTCGGAATGGAGCTCG	572
DB	21	AlaSerSerThrSerAlaSerGlySerAlaSerAlaAlaIleGlyAsnGlyApeSer	40
QY	573	GGCGCAGCAGAAAGTTCTCGAAGACCTTAATAGCAACCTTTCTACTTTGCGCG	632
DB	41	AlaAlaSerThrSerSerProGlySerAlaSerGlySerAlaSerGlySerThrLeu	59
QY	633	GGTAAGACTGTGATTTGATTTCTTTGTCGGAATTAACAACCTTCTGTGTTCGAA	692
DB	60	-----ProAs	61
QY	693	TCAATTAATTTGACGCGCAGAGGAGGACCTGAGAAAGCGGACTATCAATAGTATGA	752
DB	61	PgIuIleYLeuThrProGluGluGlyThrGluYLeuSerGlyLeuSerThrSerApy	81
QY	753	AGCTGCATGAGAGAGCCCGCAGCAGTGAATATCTGCCGAGGAGCAATATGCT	812
DB	81	AlaAlaThrGlyGlyAlaProGlySerGlyAsnAlaLeuProGluGlyGlnThrMetLe	101
QY	813	AAGCGCAAGACTCTGACGACCAATCACTCGGCTTCTTCCACAAAACCTCCAG	872
DB	101	uArgGlnAsnSerThrSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSe	121
QY	873	TGAACCTCGAATPAGCAGCAATGTGCTCTACAGTGGGCTTACCTGATGATGATTG	932
DB	121	TGluHisSerAsnSerSerAsnValSerAlaThrValGlyLeuThrGlnMetValApy	141
QY	933	TGACGACCAATGAGAGAGAAACAAATGATGTGAGGAGCAGAGAACTGTAAAGCTGC	992
DB	141	AspGluGlnSerIleValAsnIleCysSerValIlyAspGluGluAla-	157
QY	993	CCTACCAATGCTTTAAATTTAAATGATTTAGCGCTTACCTTTGATCATTTAATT	1052
DB	157	-----	157
QY	1053	GTTTTTTTTTGCTATCTTACAATTTAGTTTAACTGTAACTTGACCTTAACCTCG	1112
DB	157	-----	157
QY	1113	CGAAGCTCGATCAAAAACAGACATTTTCTTGGAACCGTAATTAAGCTCAATAAATATTA	1172
DB	157	-----	157
QY	1173	ATTGATCTTGATGAGATGATCATATGATGATGATCAAAACATCTCAAGAAAGCTCAA	1232
DB	157	-----	157
QY	1233	TTTGATCAACTAATTAGTTTGAGAAAAAATGCTGTACTTTTAAGATATATTATTA	1292
DB	157	-----	157

QY 1293 AATTTGCTGAGTAATGATATATATAGTCACAAATATTTTAGTTAAACTGCTAAAGCA 1352
 Db 157 ----- 157
 QY 1353 TTTTGAATAGCCGTGCTAGCAGAGTCTACTAGACGGGTGTAATAAGCTAATTTTATTTT 1412
 Db 157 ----- 157
 QY 1413 AAAAGCTGCTTAATATTCATTAACCATTAATGTCCTCATTTTCAGAAATTAAGTTCTAATA 1472
 Db 158 ----- Glut1leaserSerAsnly 163
 QY 1473 AGCAAAAGGTCAAGACGCTGTGGCGCTGCAAGAAAGGTTCTACATCCAGTTTGACTGT 1532
 Db 163 salalysgllyglinalalaglygllycysglunthgylserthserleuthrva 183
 QY 1533 CAAGAAGAACCCACGAGTCTTAAAGCAGTTTAAATATGAAAAAAGAAAGAAAGAA 1592
 Db 183 llysglunthpProthAspValleuglyserleuValasmetLylelyglunthrgl 203
 QY 1593 AATATCTTGCCCAACGATGCTCCCTGTGGTTTGGTTTCAATTGGTATGCAACAGACAA 1652
 Db 203 uasnHisSerProthMetSerProvalglyPheglySerileglyAsnAlaglnspas 223
 QY 1653 CTCGGCTACACCGGGTAAGTTTAAAGATCCATATTAAGCAATAACAAGAAATTAATGT 1712
 Db 223 nserAlatthr----- 226
 QY 1713 CAGTTACCAATTTTATTTAGTATGCAAAAGAACTACTATAGCGATATCTCGCTTTTAA 1772
 Db 226 ----- 226
 QY 1773 TTTTATTTTATTAAGAAATACGAATATTTCTAATTTGTAATAATAATTGATTAATA 1832
 Db 226 ----- 226
 QY 1833 CTAGAAATTTAAACCTTTTGAATTAAGACATACCTTCGAAAAATGAGTAATCATTTGG 1892
 Db 226 ----- 226
 QY 1893 AACGAGATGTGTGCCGAGAGAGACTATATAAAACCTTTGAGTATCTGATACTGCA 1952
 Db 226 ----- 226
 QY 1953 CGCTACTAAAAATGATTAAGTTAGAAAAATGGGTATATTTTGTAGSAGTTTCATTTT 2012
 Db 226 ----- 226
 QY 2013 AGAAGAAATGTATTTTATTTAAACCCCTTCAAGCGGAATCAATTTGTTCTAGATA 2072
 Db 226 ----- 226
 QY 2073 TTTTGGAAAAACAATGTTAAGTTGGAAGTCCATAAACAAGAAATTCACGGTTTCA 2132
 Db 226 ----- 226
 QY 2133 AATACTAACCAAGTTTGTGATTTTAATTTGATTAATGAGAAATTTATCACATTCAGTTA 2192
 Db 226 ----- 226
 QY 2193 AATATGTTAATTTGATTAAGTGTGACAAATCACAGCAGATTTTCATTTTTCGCGTATAT 2252
 Db 226 ----- 226
 QY 2253 ATAGAAGTCGCTTACACTCTTCTGGCGGCTTACACCTAGGTGAGATTTCGCGCGCA 2312
 Db 226 ----- 226
 QY 2313 GTGATTTATATATGATGATTTTACAGATTATTTTAATTTTATGCTGTAATTTTAATAATAT 2372
 Db 226 ----- 226

QY 2373 CTTATTTATTCATTTTACATAGTTTAAATTTGAAGAATTTCAACAGACACTACCGGAA 2432
 Db 227 ----- Pro-Vallystillegluargl1serasnspserthtrthglu 240
 QY 2433 AAAAAAGATTCGCTCTTGAACAATGAATTAATGACAAATGACATGGAGGCTGCATCAG 2492
 Db 241 lyllygllyserSerleuthrMetasnAsnAspGlnMetSerMetglunlyCyAsnGln 260
 QY 2493 TTGAATCCCGATTTTTCATATGAATTTTAAATTAATTCGCAATTTTCAGACATTAATGA 2552
 Db 261 LeuasnProaspPheilleasnglusSerleuAsnaspProAlal1serSerilleVal 280
 QY 2553 AGCGAGTAGAGACCAATATCCCGAATCCGAGATTGAGCGGGGCAAGAAATTTATGACT 2612
 Db 281 SerGlyValGlyProlethProglylileglyValGlyAlaglythrglyasnleuthr 300
 QY 2613 GCCAAGCCCAATGGAATTCCTCGGTAGCAGTAATTTGTTGATTCATGCAACAGCA 2672
 Db 301 AlaAsnAlaAsnDlyleaserSerGlySerSerAsnCysleuAspPyrmecGlnGln 320
 QY 2673 AATCACAATTTGCTGTTTCAATCACTGAGCGCCCAACAAAGGGCGCAATGATTTTAAGC 2732
 Db 321 AsnHisillePheVallePheSerthrglnleuAlaAsnlyGlyAlagluserval1leuSer 340
 QY 2733 GGTCAATTTTCAACTATTTATGCGTATCACTGCACTCAGCTCTCAAAAAAGCTTCGT 2792
 Db 341 GlyGlnPheGlnthrllel1leAlatyrh1scyThthGlnProAlathrllysserPheleu 360
 QY 2793 GAAAGCTTTTATTTGAATAAACCTTTAAAGATTAACAAGTTACAGCGGCACAATCCGTC 2852
 Db 361 GluAspPheMetCysAsnProleuLyAlleAsnlyleuGlnAlrgh1saAsnSerVal 380
 QY 2853 GGTATGCCATGATTAAGCAGTGGGGAGTTGACCTAATCCCTATCTGTAGCCAAA 2912
 Db 381 GlyMetProthrl1lelyMetGlyGlnValglyleuthrProAsnProvalAlaLys 400
 QY 2913 ATAACAACAAGCAGCAGCACAATCAAAAGACCGTAGGCTTATTTGAACCCCAATTCAA 2972
 Db 401 l1ethrGlnGlnInthProh1sthrlysthrValGlyleuLeuLysProGlnPheAsnGln 420
 QY 2973 CATGAAAAACGAAACGTAAGTACTGTAAGCGCGCTTACAGCACTTTTGTGACAGCTT 3032
 Db 421 HisGluAsnSerlysaSerthrlVal1serAlaProSerAsnSerPheVal1AspGlnSer 440
 QY 3033 GATCCTATGGGCAACGAACTGAATTCGATGAGCGCGCTTACAGCACTTTTGTGACAGCTT 3092
 Db 441 AspProMetGlyAsnDlythrglnleuMetCysTrpGlnlyGlySerSerAsnth1Ser 460
 QY 3093 AGGCTGGAACAACACTCACGAATCATGTAGACAGTATCATTCACAGCAGTCACAG 3152
 Db 461 ArgSerGlyGlnAsnSerArgAsnHisvalAspSerilleSerthSerSerGlnSerGln 480
 QY 3153 GCAATTAAGATCTGGAAGCAGCTGCTGATTTTGGACAGGTCACAAAAGAACGAT 3212
 Db 481 Ala1lelyse1leuGlnAlaAlaglyValAspLeuglyGlnValThrllysglySerAsp 500
 QY 3213 CCTGGCTGACAACTGAAAAACAATTTATTCATCTCAGAGGATTAAGTTTTCAGAGAA 3272
 Db 501 ProGlyleuthrthrglnAsnAsn1levalSerleuGlnDlyVallyllyserValProAspGln 520
 QY 3273 AACCTTACACACAGCAGCGCAACATCGGAGAAACAGTTGGCAAAAATTAATAAATATG 3332
 Db 521 AsnleuthrProGlnthrglnAsn1h1sArgGlnGlnleuAlaLys1lelyblyMet 540
 QY 3333 AATCAATTTCTTTTCTGAAAAATGAGAAATTCAGTACAGCTATGTAAGCTTACCTCACATA 3392
 Db 541 AsnGlnPheleuthrProGlnAsnGlnAsnSerValGlyAlaAsnVal1SerSerGln1le 560
 QY 3393 ACAAAAATTCACAGAAATTTAATGATGGGAGTGTGGGTGGCGGAGCGGATTCATTTATA 3452
 Db 561 ThrLys1leProDlyAspLeuMetGlyMetSerGlyGlyGlyGlySer1le1le 580
 QY 3453 AATCCAGATGAGAACATGCAATATGACAGTAAAGCCAAATCGAGCTTTATCGCG 3512

Db 581 AsnProthrMetArgInLeuHisMetProGlyAsnAlaIysSerGluLeuLeuSerAla 600
 QY 3513 ACNAGTTCCAGACTTTCCGAGAGATGTAATGATCCAGGGGATGTTATATCAGATATGGGT 3572
 Db 601 ThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMetGly 620
 QY 3573 GCCGTAATAGATGTAATATAATATCAAAAAACCAAGTGTGCATGTGATCTGAGTAGTGGT 3632
 Db 621 AlaValIleGlyCysAsnAsnAsnGlnIysThrSerValGlnCysGlySerGlyValGly 640
 QY 3633 GTTGTCTACGTGAAACACTGCACTGAGTGAATAGTCAATATGATTTGCTCAAGCTCCGGC 3692
 Db 641 ValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGly 660
 QY 3693 GCCCCCAATGGCAATATGATGGAGAGCTCTACGATATGCTAGCTCGTTTGGCAACACA 3752
 Db 661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
 QY 3753 AGCTGCAACGCTCATCGGAACGGCCAGATATGTCTTAAGGAAGTTTAAATCAAGATAC 3812
 Db 681 SerCysAsnValIleGlyThrAlaProAspMetSerIleGlyValLeuAsnGlnAspSer 700
 QY 3813 CGAACCCATTCACATCAAGGGGAGTTGCTCAATAGAGTGTGGAAGATTCAACATCA 3872
 Db 701 ArgThrHisSerHisGlnGlyValAlaGlnMetGluTrpSerIysIleGlnHisGln 720
 QY 3873 TTTTTCGAAGAAGCGCTCAAGGGGGCAAGCCAGACAAATGCTGGAATCTGTAATGACCA 3932
 Db 721 PhePheGluGluArgLeuLeuGlyValAspProArgGlnValThrGlyThrValValPro 740
 QY 3933 CAACAGCAAAACCCCTTCTGATCTGTGTGGAACCTCGTTAAACACAGGTGCGACCCCTG 3992
 Db 741 GlnGlnGlnThrProSerGlySerGlyValAsnSerLeuAsnGlnValAlaGProLeu 760
 QY 3993 CAAGGTCCACCTCTCTTACCACTGCATCCAGAAAGATCTGCGTCAATGCAATAGCCACT 4052
 Db 761 GlnGlyProProProProGlyHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
 QY 4053 CAATGCCCAATCCCTCGAGTCCCAACAATCTATCTCCCGTCACGCGGACCAACCGCA 4112
 Db 781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThrAla 800
 QY 4113 GCAGTCATGGAGATTGCGCAACAATCTCTAGCATGATGATGAAACAGATCATTAATGCA 4172
 Db 801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
 QY 4173 TCTGTTCCCGAAGCTAATATCTTGAGAGGTTGAGGAGGACACAAACAGTCTCTACGA 4232
 Db 821 SerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrValLeuSerAla 840
 QY 4233 AACCAAACTGTTTTCAGGAGACACCCCATCGCGTCAAAATCAAAATCGTAGTAGAAAT 4292
 Db 841 AsnLeuAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAsnArgSerHisGln 860
 QY 4293 ACCGATCGTCAACGCTTCTTACGATTAATTAAGACAGCAACCAAGTACCCCTTATCT 4352
 Db 861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
 QY 4353 CATCTATCCCAAGAAATTTGAGTCTTTCCGTCATCTCTGCTGCTGATATATATTTG 4412
 Db 881 HisLeuSerProGlyGluPheGlnSerPheGlyGlnSerSer 894
 QY 4413 TTTAATTTTTTAAAGCAAAATCAAAATATGAAATTCGCTTAATATAATATATTACA 4472
 Db 894 894
 QY 4473 TAACTCGAAATTTGATAGAAAAATCAGGAATAGAAAAATTAATATTTTCCGAGACG 4532
 Db 894 894
 QY 4533 CCATCATTTTCTTGAATCCAAATTTCTGAGTAGTATTGATAGATATCTACTATTAAAA 4592

Db 894 894
 QY 4593 TTAACACAGAAATTCATCCGTTAATGAAATCACTATTGTTAATAGAAATTAAA 4652
 Db 894 894
 QY 4653 AATATGTTTATTATATATATTTCTACAGGTGATTAACAGAAAAGTAGGCGAACGCCAC 4712
 Db 895 895
 QY 4713 AGGTCAGGGGTCAACAGTAATTAATGCTTAATAGAGCAATTAAGTATACATTTGCTG 4772
 Db 907 GlnGlnArgSerProValAsnSerLeuIleGlnAlaAsnValAspValArgPheAla 927
 QY 4773 CATCCAGTCTGTTTAAACCCGATCCCATATGCAAGCAATTCGAATTCAGACTTAA 4832
 Db 927 IeSerSerProGlyPheAsnProHisProHisMetGlnSerHisSerAsnSerAlaLeu 947
 QY 4833 ACGCTATTAATATGGGCTCTACCAATATACAGATGGAGGTAATATTAAATATTTTAT 4892
 Db 947 snAlaIyIysMetGlySerThrAsnIleGlnMetGlu-Arg- 960
 QY 4893 TAAAGTTTGTGTTAATTTATCTTTTTCAGCCTCAAGCATCAGCGCAAGTGGATC 4952
 Db 961 961
 QY 4953 CGTACAAATTAATCGCGCTCCGATATATTCGCTAATTCCTCAATAGTGCATATGCGC 5012
 Db 968 ValGlnPheSerArgArgSerHisAsnIleProLeuAsnProAsnSerGlyAsnArgPr 988
 QY 5013 GCCACCAACAAAGATGACCCCAAACTTGATCCATCTCTTTGGCAACAAATGTCCA 5072
 Db 988 OProProAsnIysMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerG 1008
 QY 5073 ACAACTAACAAATTTGGGTCTCCAGATGGGTGTCCAGCGGCAACGTGATGATGACAT 5132
 Db 1008 nGlnLeuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGlyMetThr 1028
 QY 5133 GATGGGGGGTCCGGGACGCTCCGACATCAATATTGAGCATGGAATATTCCGAGCTAGA 5192
 Db 1028 tMetGlyGlyProGlyProSerAspIleAsnIleGlnValGlyIleIleSerGlyLeuAs 1048
 QY 5193 TGGATCAGGAATGATATCCATTAATCAAAATTAATCTGTCATTCATGATGCTTAATGA 5252
 Db 1048 pGlySerGlyIleAspThrIleAsnGlnAsnAsnCysHisIleSerMetAsnValAlaMetAs 1068
 QY 5253 CTCATATGGGTCCCGGATGCTGAATCTTAAATGCGGTAAGACGGCGTCCAAATGGAC 5312
 Db 1068 nSerMetGlyProArgMetLeuAsnProIysMetCysValAlaGlyGlyProAsnGlyPr 1088
 QY 5313 GCTGGCTTTAATCCTAATTTCCCAATGATGATTAAGAGAAATTCATAGGCTCTG 5372
 Db 1088 OProGlyPheAsnProAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerG 1108
 QY 5373 CTGTGGCTAGCAAACTTTTCAAACTTTCAAGGGGTTGTTCCACTGTGTGCGCAATGAT 5432
 Db 1108 YCysGlySerAlaAsnSerSerAsnPheGlnGlyValAlaProProGlyAlaArgMet 1128
 QY 5433 GGGTGAATGCGAGTCAATTTTGGTTCGAATTTCAATCCGATATTCAGTTAAAGCGAG 5492
 Db 1128 tGlyArgMetProValAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSe 1148
 QY 5493 TACCCCAACACCATATCAATATACATGCAATGAGGACAGCAACGCCCAACAAATTAACA 5552
 Db 1148 tThrProAsnThrIleGlnIlytMetProValArgAlaGlnAsnAlaAsnAsnAsnAsn 1168
 QY 5553 CAATGAGCTAATATATGTCGAATGCGACCTTGTCTGGAATTTTTCAGAGATGCTTAA 5612
 Db 1168 nAsnGlyAlaAsnAsnValArgMetProProSerLeuGlnPheLeuGlnArgTyrAlaAs 1188
 QY 5613 CCTCAAAATGGGTGCTGTAGGCAATGGGTGCGCAATATGCGCAATGAGCGAGCGAG 5672
 Db 1188 nProGlnMetGlyAlaValAlaGlyAsnGlySerProIleCysProProSerAlaSerAspG 1208

QY 5673 TACTCTGGAATCCAGATTTGATGGGAGCAGAGCCGAGTATGTAATGATTC 5732
 1208 ythrProglYmeProglYleuMetAlaGlYProglYAlaGlYglYleuMetLeuMetLense 1228
 QY 5733 TTCGGAGAGCAACACCAAGAACATCCTGGGCGCAAGCATGGTATTA 5792
 1228 rserglYglYlnhIsglYlnhIserIleThraAsnProglYAlaSerAsnIlyIleA 1248
 QY 5793 CTTCCTTGAATTCGAATCAATGCTATGTTGAGCAAGAGGGTGGATTACCGGCCA 5852
 1248 nPhePheGlnAsnCysAsnGlnMetSerIleValaPglYglYglYleuProglYhI 1268
 QY 5853 TGAAGATCAATGAATTTGGTCAACATCTATGATTAAGGGGATGGGTCACATGGCAT 5912
 1268 sAaPglYserMetAsnIleGlYglYlnProSerMetIleAglYglYleuArgProhIsAlaMe 1288
 QY 5913 GCGGCCAATGTAATGGTGGCGGAGTCCAGCCGTTTAAAGCAATTCAGTTTGACA 5972
 1288 tArgProAsnValMetGlYAlaArgMetProProValaAsnArgGlnIleGlnPheAlaGl 1308
 QY 5973 GTCATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6032
 1308 nSerSerAspGlyIleAspCysValGlYAspProSerSerPhePheThraAsnIleSerCy 1328
 QY 6033 CAACAGGCGTGGACACACATGTTGGATGAGCAACAGGCGCAATCAGCCTTAAGCACA 6092
 1328 sAaSerAlaGlYProhIsMetPheGlnSerIleAlaGlnIleAlaAsnIleProlyThrGl 1348
 QY 6093 ACACATAAAGAACATACCTAGTGAATGTGTCAAAACCAATCGGAGCTTGACAGTGACACA 6152
 1348 nhIstIleYAsnIleProSerGlYMetCysGlnAsnGlnSerGlYleuAlaValaIleGl 1368
 QY 6153 AGGCGCATGCAACCTGATGGGCAAGACATGCGGCGGTCAGTCTTTAATGGACCTAC 6212
 1368 ngIYglnIleGlnLeuIleIsglYngIYnhIsAlaGlnIYglYlnSerleuIleGlYProTh 1388
 QY 6213 TAATATTAATTAATGATCACTGCGGAGTGTCAAGTCACTTAACGAGTCTTGGCAT 6272
 1388 ThraAsnAsnLeuMetSerThraIleGlnSerValSerAlaThraAsnIlyAlaSerGlYI 1408
 QY 6273 CAATTCGTAGTCCCTCTTCTTACGAGCTGAAGTATGCCACCAATATCATGATTTC 6332
 1408 eAaPheValGlYProSerSerThraPheLeuYsYrAlaGlnIlnhIsSerPheGl 1428
 QY 6333 GCAAGCATTTATGATCAACACAGAGTCAACAAACAGCATATGACACGACGCA 6392
 1428 ngIngIlnLeuYrAlaThraAsnThraArgSerGlnGlnIlnhIsMetIleGlnIlnhI 1448
 QY 6393 CCAAGAGCAATGATCAATGCGCGGATTTATCACAAATCCCAACGTTCTTGTGCA 6452
 1448 sGlnSerAsnMetIleThraMetProAsnLeuSerProAsnProThraPhePheValAs 1468
 QY 6453 CAAA 6456
 1468 nIlyS 1469
 RESULT 2
 Q67FY3 PRELIMINARY; PRT; 1530 AA.
 ID 067FY3
 AC 067FY3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Bcl9-2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_taxonomy=7955;
 RX 1)
 RP SEQUENCE FROM N.A.
 RA Brembeck F.H., Schwarz-Romond T., Bakkers J., Wilhelm S.,

RA Hamerschmidt M., Bircheimer W.;
 RT "Essential role of BCL9-2 in the switch between [beta]-catenin's
 RT adhesive and transcriptional functions."
 RL Gene Dev. 18:0-0(2004)
 DR EMBL; AY296057; AA062695.1; -
 SQ SEQUENCE 1530 AA; 159872 MW; C29FEC9433ED28C0 CRC64;
 Alignment Scores:
 Pred. No.: 9, 93e-14 Length: 1530
 Score: 338.50 Matches: 308
 Percent Similarity: 30.82% Conservative: 171
 Best Local Similarity: 19.82% Mismatches: 554
 Query Match: 2.81% Indels: 522
 DB: 2 Gaps: 71
 US-10-664-859-1 (1-6909) x Q67FY3 (1-1530)
 QY 2565 CCAATACCCGGAATTCGAGTGGAGCGGAGCAAGAAATTTATGACTGCCAACGCCAAT 2624
 153 ProserGlnIlyhIsSerGlYserSerThrThrIlyhIsValSerAspProglYglYPro 172
 QY 2625 GGAATCTCTCGGGTACGAGT-----AATGTTGGATTAATCAACAAAGCA 2672
 173 GlYleuGlYserGlYhIsGlYProglYIleArgThraPheLeuIleSerArgProglYn 192
 QY 2673 AATCAATATTCGTTGTTTCACTGAGTGGCCCAAAAGGGGCGAATCAGTTTAAAGC 2732
 193 GlnValIlyAlYValPheThrThrSerIleAlaSerIleAlaIleGlnAlaValMetIle 212
 QY 2733 GGTCAATTTCAACATTAATGCGTATCACTGCACCTGACCTGCTCAAAAGCTTCTTG 2792
 213 GlYhIsThraPheSerIleLeuLeuYrnhIs--GlnGlnAsnValProArgThrIlyLeu 231
 QY 2793 GAAAGCTTTTATGAAAAACCTTAAAGATTAAAGATTAAAGGGGCAATTCCTGC 2852
 232 AspGln-----SerThrGlYAlaGlYIlyValSerAsnLeuIleGlnIleSerSer 249
 QY 2853 GGTATCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2912
 250 SerhIs-----SerProProIleGlYhIProIlyS 259
 QY 2913 ATTAACAAACAGAGCCACATACAAAGCCGTAGGCTTATGAAACCCCAATTCATCA 2972
 260 SerGlnSerGlYThrProArgProIleSerValGlYAla----- 273
 QY 2973 CATGAAACAGCAAAAGCTGATCTGTAAGCGGCTTACAGCACTTTTGTGACAGTCT 3032
 274 -----GlYhIsLeuProglYThrSerThraPheSerThrGlYhIsProAspSer 290
 QY 3033 GATCTATGGGCAAGAACTGAATGATGCTGGGAAGCGGATCTCAACACACAGT 3092
 291 GlnProAlaGlnThrhIs-----ArgGlYglYglYThr----- 301
 QY 3093 AGGTCTGACAAAACTCAGGAATCATGTAGACAGTATCATGCA--TCCAGCGAGTCA 3149
 302 ---SerSerAsnAsnGlYArgSerAlaValhIsThrleuGlYleuGlYAsnSerGlYPro 320
 QY 3150 CAGGCAATTAAGATGATCTGAGACAGCTGGCGTTGATTTGGGACAGGTCAAAAGCAAGC 3209
 321 GlnSerValGlYValSerGlYThrGlYglYValaAspArgProglYAlaIleProhIsIle 340
 QY 3210 GATCTGGCGCTGCAACATGAAAACATTTGATCACTGCAAGAGATGATGTTCCAGAC 3269
 341 GlYAlaGlYValSerProSerThraPheSerProSerValleuSerAlaIleuArgGlnSerGln 360
 QY 3270 -----GAAACCTTACACCAACCAAGCGGCAACAT 3299
 361 leuGlYglYAlaGlYProglYAsnThrAspGlYleuSerIlyglYlnhIsIleuGlnhIs 380
 QY 3300 CCGGAAGAACAGTGGCAAAATTAATAATCAATTTCTTTCTGAAAATGAG 3359
 381 ArgGlYArgSerleuGlnIlnhIleuArgAspIleGlYArgleuLeuArgSer----- 398

QY 3360 AATTCAGTAGAGCTAATGTAAGCTCAGAGATAACAATAATTCAGAGATTTAATGATG 3419
DB 399 -----GlyThrGlyValAlaGlnGluAspProArg----- 408
QY 3420 GGGATGTCGGGTGGCGGAGCGGATCTATTATAATCCGACGATGGACAACCTGCATATG 3479
DB 409 GlyProAsnGlyAsnProAsnGlyThrAsnValAsn----- 420
QY 3480 CCAGTAAGCCCAATCGAGCTTATTTCGGCGACAAGTTCAAGACTTTCGAGAAGATGA 3539
DB 420 ----- 420
QY 3540 ATGCATCCAGGGAGTGTATATCAGATATGGTCCGTAATAGAGTAAATAATCA 3599
DB 421 -----AsnAsnAsnSer 424
QY 3600 AAAACGAGTGTCAATGTGATCTGGAGTAGTGTTGTCACTGGAACAACCTGCAGCTGA 3659
DB 425 Asn-----AspGlyGlyArgGlyLeuGluAspGlyGlyLeuGlyGly 439
QY 3660 GTAATATGCAATATGATGCTTGCAGCTCCGCGCGCCGCAATGCAATATGATGGAGAG 3719
DB 440 IleProGlyAsnGlyAsnIleAsnAsnAlaGlyMetProGlyMetProProValGlyGly 459
QY 3720 TCT-----ACGATATGCTAGCCTCGTTGGCAACAAGCTGCAACGTCAATCGA 3770
DB 460 MetLeuLeuSerGlyGluProLeuGlnSerIleIleSerGlnThrGlnAsnLeuGlyGly 479
QY 3771 ACGGCCCAAGATATGCTAAGAAAGTTTAAATCAAGATAGCCGAACCCATCACTCA 3830
DB 480 ProGlyLeuAspAspSerLeuMetGlyProIleHisGlyMetProProHisSerHis 499
QY 3831 -----GGGGAAGTTGCT--- 3842
DB 500 LeuSerSerProSerGlyLeuAspMetGlyProLeuLeuGlyProGluGlyValThrPro 519
QY 3843 ---CAATGAGAGTGTGCGAAGATTCAACATCAATTTTTCGAAGA---CCCTCAAG 3893
DB 520 GlnGlnLeuAlaThrArgGlyLeuGlnGlnGlyTyrGlnGlnGlyAspArgGlnHis 539
QY 3894 GGGGGAAGCCGACAGTCAAGTCACTGGA-----ACTGATACCAACAAGCAACCCCT 3947
DB 540 AspMetAsnProIleValGlnHisProGlnHisPheArgIleMetProGluMetGlyMetPro 559
QY 3948 TCTGATCTGTGGTGAACCTGTTAAACAACAGGTGCGACCC---CTGCAAGGTCCA 4001
DB 560 GlyGlyPro-----PrometLeuMetArgGlyPro 569
QY 4002 CCTCTCTCTTACCACTCC-----ATCCAGAGATCTGCGTCAAGTACCAATAGCC--- 4049
DB 570 ProProProTyrHisSerIleAspProGlyAspGlnGlnThrGlyProGlyPrometValGly 589
QY 4050 -----ACTCAATCGGCCCAATCCCTCGAAGTCCA 4076
DB 590 GlyGlyMetGlyGlyAsnAlaArgMetMetAspMetAsnGlnGlnGlyProArgGlyPro 609
QY 4077 AACCATCTATCTCTC-----CCGTCA-----CCGCGACACAACGCGA 4112
DB 610 ArgPheLeuGlyGlyMetArgGlyProSerGlyGlyGlyGlyTyrProGluSerProGly 629
QY 4113 GCAGTCATGGGA----- 4124
DB 630 GlyValLeuGlyValGlnGlyLeuGlyProGlnArgProProArgProGlyMetGlyTyr 649
QY 4125 TTGCGCGACAACCTCTAGCATGAGTAGAAGAGATCATTAATCTGATGTTTCCGCA 4184
DB 650 LeuGlnGlnIleProProAsnMetGlyGlyGlyProPheHisGlyCysCysProPro 669
QY 4185 GCTAATACTTCAGCGGTTCAAGGACGACGAACAACAGTCTTCAGCAACAAGAACTGT 4244
DB 670 GlyGlyPro-----GlyGlyProProGlnHis 678
QY 4245 TTTCAGGACGACACCCCATCGCCGTCAAT----- 4274

DB 679 PheGlnGlyAspLeuAspArgPrometThrArgGlnGlnIleTyrArgArgIleHisArg 698
QY 4275 -----CAAAATGTAGTGAATAACCGAGTGTCAAGGCTCTTACGATACATACTA 4325
DB 699 LeuAspLeuGlnGlnMetSerArgGlnGlnGlnGlnGlnGlyLeuGlyGlyProArgMet 718
QY 4326 AGCAGCAACCCCAAGTACCCCTTATCTATCTATCCCAAGAAATTTGATGCTTTCGGT 4385
DB 719 MetAspAsnProGlyGlyPro-----GlyPheProAsnProGlyMetAlaGlyGly 725
QY 4386 CAGTCTCTGCTGGTATGTTATATTTGTTAATTTTAAAGACAAATCAATATGAAT 4445
DB 736 ProProSerArgGly----- 740
QY 4446 TGCGTTAATAATAGTTATATATATACATAACTCGAAATTTGATAGAAAATCAGAAAT 4505
DB 741 -----AspProMetAspPhe 745
QY 4506 AGAAAAATTAATTAATTTTCCGACCGCCCATCTATTCTGAATCCAAATTTCTGAGATG 4565
DB 746 ProValSerArgThrIleMetGlySerProIleGlyGlyValGlyGlyAspGlyGlyPro 765
QY 4566 ATTTGTAAGATTAATCTATTAATAATTAACAGAAATTCATATCCGTTAATTGAAA 4625
DB 766 ThrMetArgAspIleValAspSerProLeuGlyGlyAsnLeuAsnMet----- 781
QY 4626 ATCACTATGTTTAATAAGAAATTAATAATGTTATATATATTTCTCAAGTGATAT 4685
DB 782 -----AsnMetGlyMet--AsnIleAsnGlnGlnGln 792
QY 4686 ACATGAAGAATAGGACGACCAAGCCCAAGGTCACGCGTCAACAGTAATAGTCTAATAG 4745
DB 792 IleLeuAlaGlnLys-----LeuArgGlyGlyProGlyValLeuGlyGlnMetLeu 810
QY 4746 AGCAATAATAAGATGACATTTGCTGCATCCAGTCTCGTTTAACCGCATCCACATA 4805
DB 810 AsnAlaGluAspIleSerArgGlyLeuArgAlaSerGlnAsnGly----- 823
QY 4806 TCGAAGCAATTCAAATTCAGCATTAACGCTTAATAATGGGCTTCAACAATATACAG 4865
DB 824 -----ArgGlyGlyAlaAsnLysAlaHis 831
QY 4866 TGGA-----GGTAAATATTAATAATTTAATTTAACGTTTGTGTTAATTT 4912
DB 831 eIleProGlyProGluGlyPro----- 838
QY 4913 ATCTTCTTTTCAAGGTCAAGCATCGCGCAAGTGTGATCCGA----- 4956
DB 839 --LeuGlnPheProAsnGlnSerPheProGlyGlyGlnValAspGlyProTyrLeuG 858
QY 4957 -----CAATTTA 4963
DB 858 GlnGlnProGlyProAspMetTyrGlyProAspGlnProGlyProProHisLeuSerSer 878
QY 4964 GTCCGGCTCCGATATATATTCGCTAATCCAAATAGT-----GGCAATC 5008
DB 878 HisSerArgLeuSerHisIlePrometAsnThrGlySerArgValThrAspLeuGlyAla 898
QY 5009 GGCCTGACCAACAAGATGACCCCAAACTTGATTCATCTCTTTTGGCA---CAA 5065
DB 898 ArgHisProProAsp---LeuProIleSerValAsnPrometGlySerProAlaIlePro 917
QY 5066 TGTCCAACAACCTCAACAAGTTGGTGTCCAGC-----ATGGTAGTAC 5107
DB 917 roSerHisGlnLeuLysSerProSerLeuSerGlnLysProSerProLeuMetProSerP 937
QY 5108 CA-----GCCGAACCTGTGTATGACGATAGGAGGGGTCCGGGACCGTCCGACATCA 5161
DB 937 roSerAlaAlaGlyLeuLysSerProSerGlnLeuProGlnSerGlyProThr----- 954
QY 5162 ATATTGACATGGAATATTTTGGGACTAGATGATCAGAGATAGTATCATTAATCAAA 5221

QY 2475 ATGAGAGCTGCAATGATGATCCCGATTATTCATGATCTTTAAATAATCTCGCA 2534
 Db 208 ThrAsp1LeThrMetAsp1LeuMetValLeuMetLeuThrAla1aThrSerProSer 227
 QY 2535 ATTTCAGCATATTAGTAAAGCGGAGTAGCA-----CCAAATCCCGGAATCGCA 2582
 Db 228 ThrAlaVal1ThrSerProSerSerLeuGly1ThrSerSerSerProLeuProSer----- 245
 QY 2583 GTTGAGCGGGGAGGAAATTATGACTGCGCCCAATGGAATCTTCGGGTAGC 2642
 Db 246 ---Ser1LeSerThrSerAlaLeuPro1Lea1aSerSerSerAlaSerSerProSer 264
 QY 2643 AGTAATGTTTGGATTACATGCAACAGCAAAATCACATATTCGTGTTTCACTGACGTG 2702
 Db 265 Ala-----AlaSerSerThrThrProValValLeuSerSerThr1Le 279
 QY 2703 GCCAACAAGGGGCGCAATCAGTTTAAAGCGGTCAATTTCAAACTTATTTGGCTATCAC 2762
 Db 280 GlnSerSer-----SerGly1ThrPheProSerSerValAla----- 291
 QY 2763 TGCACACGCTGCTACAAAAGACTTCTCGAAGACTTTTATGAAAAACCTTTAAAG 2822
 Db 292 ---SerSerProSerThrValGly-----Ser 299
 QY 2823 ATTACAGATTACAGCGGCAATTCCTCGTATGCAATGATGAGCATGCGGCGAGGT 2882
 Db 300 ThrSerGlyAla1a1aSerSerSerSerTyAla1aThrValSerThr1Lea1aGlySerThr 319
 QY 2883 GGACTACTCTCTCTAATCTCTGAGCCAAATTAACACACAGACGACCATACAAAGACC 2942
 Db 320 GlySerThr-----1LeThrProValProGlySerSerSerThr 332
 QY 2943 GTAGCGCTATTGAACCCCAATTCATCAACATGAAGAAAGCAAGAGTACTGTATGAGC 3002
 Db 333 1LeGlySerSerThrProSerAlaSerSerSerSerSerGly1ThrMetSerThr1LeSer 352
 QY 3003 GCGCCTAGCACTCTTTTGTGCAC-----CAGTCT 3032
 Db 353 GlySerThrGlySerThrVal1aThrVal1aProGlySerSerSerThrPheAlaSerSer 372
 QY 3033 GATCTATGCGGCAACGAACCT-----GAATTGATGCTCGGAGAGCGGATCTCA 3083
 Db 373 ThrPro1Lea1aSerSerSerSerProGlySerThrValThrVal1a1aProGlySerSer 392
 QY 3084 AACAC-----AGTAGCTGTGACAAATCACTGCAAGAAATCATGTAGACAGATATC 3131
 Db 393 SerThrTyGlySerSerThrProSerAlaSerSerSerSerSerGly1ThrMetSerThr 412
 QY 3132 AGTACATCCAGCGAGTCACAGCAATTAAGATACGGAAGACGCTGCGTGAATTTGGGA 3191
 Db 413 AsnSerGlySerThrGlySerThrVal1aThrVal1aProValSerSerSerThrPheGly 432
 QY 3192 CAGGTACA-----AAAGAAAGCATCTGCGCTGACCACTGAAAAACAC 3236
 Db 433 SerSerThrPro1Lea1aSerSerSerSerSerGlySerThrVal1aVal1aSerGly 452
 QY 3237 ATGTATCACTGCAAGAGATTAGTTTCCAGAGAAACCTTACACCAACAAGCGGCA 3296
 Db 453 SerSerSerThrTyGlySerSerThrProSerAlaSer-----SerSerSerAla 469
 QY 3297 CATCGGAAGAAACATGCGCAAAATAAATAAATGATCAATTTCTTTTCTCGAAAT 3356
 Db 470 GlyThrAlaSerThr1LeSerGlySerThrGlySerThrAlaThr1LeVal1aProGlySer 489
 QY 3357 GAGAAATTCAGTAGAGCTAATGTAAGTCAACAGATAAATAAATTCAGAGAGATTAAAG 3416
 Db 490 SerSerSerValGlySerSerThrGlnSerAlaSerProSerSerProGly---ThrMet 508
 QY 3417 ATGCGGATGTCGGTGGCGGAGCATTTATTAATCCAGCATGCGACAACTGCAT 3476
 Db 509 SerThrValSerGlyProThrGlySerThrVal-----ThrVal 521

QY 3477 ATGCCAGTAGACCCCAATCGAGAGCTCTTATCGGCGCAAGATTACAGACTTTCGAAGAT 3536
 Db 522 ValProGly1SerSerThrSerPro1aProSerSerSerPro----- 535
 QY 3537 GTAATGCATCCAGAGGATGTTATATCAATATGGGTCGGTATATAGATGTAATAAT 3596
 Db 536 -----AsnProSerSerSerPro1aSerThrGlySerThr1LeThr1LeSerGlySer 553
 QY 3597 CAAAAAACAGATGCAATGTGATCTGAGTAGAGTGTGCACTGACCAACATCGACGT 3656
 Db 554 SerSer1Le1LeValSerThrValSerGly-----SerThrValSerGlySerThr----- 570
 QY 3657 GAGTAATATGTCATATGATGATGCTCAAGCTCCGGGCGCCG---AATGCAATATGATG 3713
 Db 571 GlyThrSerGlnSerThr1Leu1a1aSerSerThrAla1aThrProGlySerSerSerThrVal 590
 QY 3714 GGAAGCTTACGATATGCTAGCTCGTTGGCAACAAGATGCAAGCTCATCGGAAGC 3773
 Db 591 ProSerSerSerSerProGlnProSerSerGlnSerPro1aProAsnThrGlySerThr 610
 QY 3774 GCCCAGATATGCTTAAGGA-----GTTTAATCAAGATAGCCGAAC 3818
 Db 611 ThrProSerGlnThrSerSerGlnSerProSerProSerMetAsnProSerSerThr 630
 QY 3819 CATTCACATCAAGGGAGAGTCTCAATAGAGTGTCGAAGATTCAACATCAATTTTC 3878
 Db 631 ProThr-----GlySerSerGlnSerThr1LeThr 640
 QY 3879 GAAGAAAGCTCAAGAGGGGCAAGCCCAACAAGTACTGGAAT----- 3923
 Db 641 ProGlnGlySerThrAlaSerSerProThrGlySerThrGlySerThrPheSerValAla 660
 QY 3924 -----GTAGTACCAACAAGCAAAACCCCTTCGATCT-----GATGAAATCGTTA 3971
 Db 661 ThrGlnVal1aThrSerGlnSerThrValProSerGlySerSerLeuGly1ThrGlnSerThr 680
 QY 3972 AACACCAAGTGCGAGCCCTCGAAGGTCCACTCTTACACTTCACAGAGATCT 4031
 Db 681 AsnSerSer-----ProSerProSerSerSerSerProSerThr 693
 QY 4032 GCGTACGTACCAATAGCCCACTCAATGCCAATCCCTGAGTCCCAACAACATATCTTC 4091
 Db 694 SerGlyMetSerThr1LeuThrSerGlnProSerProSerSerThrGlnSerSerGlyAla 713
 QY 4092 CCGTACCGCGCAACAACCGCACAGATCATGAGATTGCGCAACCACTCTCTAGCATGAT 4151
 Db 714 GlnSerThrLeuThr-----ProSerProAsnProSer-----Gln 726
 QY 4152 GGAACAGATCATTTATCTGATCTGTCCGCAAGCTAATATCTTCAGCGTTACGCAAGC 4211
 Db 727 SerThrSerSerLeuGlnSerSerThrSerGlyAla1aThrThrSerSerGlySerAlaGly 746
 QY 4212 ACAACAACAGTCTCTCACAAACAAGAACTGTTTCAGGACAGACCCCATCGCGTCA 4271
 Db 747 ThrThrMetThr-----SerProSer 753
 QY 4272 AATCAAAATCTAGTAGAATAACCGGATCGTCAAGGCTTTTACGATAAATTAACTACAC 4331
 Db 754 -----GlnSerSerSerValGlySerSerGlnGlySerThrSerPro1a1aSer 770
 QY 4332 AACCCAGT-----ACCCCTTATCTCATCTATCC 4361
 Db 771 ThrThrSerGlyGlnMetThrSerGlnGlySerThrGlnThrProGlySerSerValSer 790
 QY 4362 CCAGAGAAATTTGAGCTTTCGTCAGTCTCTGCTGCGATATGTAATTTGTTAATTT 4421
 Db 791 -----ThrSerAla1a1e1e1e----- 796
 QY 4422 TTTAAAGCAATCAATATGAATGCGTAAATTAATATATATATACATAACTCGCA 4481
 Db 797 -----ThrSerThrGlnGlnSerValSerThrAsnSer-----ProGly 809
 QY 4482 AATTGATAGAAAAATCAGAAATAGAAAAAATAATTAATTTTCCGACCGCCCATTCAT 4541